

**Estudio de la influencia política de la red social X a través del análisis de sentimientos y  
clasificación de textos**

Edwin Alberto Mendoza Moreno

Sebastian Fernández Barreto

Asesor

Isaac Esteban Camargo Freile

Universidad Nacional Abierta y a Distancia UNAD  
Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI  
Especialización en Ciencia de Datos y Analítica

2025

### **Dedicatoria**

Este trabajo está dedicado con cariño a nuestras familias, quienes con su amor, apoyo y paciencia nos han acompañado en cada etapa de este proceso. Finalmente, queremos dedicar esta tesis a todas las personas que creen en el poder transformador de la investigación y la tecnología. Esperamos que este trabajo aporte al entendimiento de la influencia de las redes sociales en la política y su impacto en la vida cotidiana.

## **Agradecimientos**

Queremos expresar nuestro más sincero agradecimiento a nuestro director de monografía, cuya sabiduría, orientación y dedicación han sido esenciales para el desarrollo de esta investigación.

Gracias a su guía, aprendimos a enfrentar cada desafío con disciplina, perseverancia y una profunda curiosidad intelectual.

A nuestras familias, por ser nuestra mayor fuente de motivación y apoyo incondicional. Gracias por su paciencia, por brindarnos su amor inquebrantable y por confiar en nosotros en todo momento. Sin su apoyo, este trabajo no habría sido posible.

Extendemos nuestro agradecimiento a todas las personas que, de alguna manera, contribuyeron a que esta monografía se hiciera realidad. Cada conversación, cada palabra de aliento y cada gesto de apoyo han sido fundamentales a lo largo de este proceso.

Por último, damos gracias a la vida por las oportunidades, los aprendizajes y los desafíos que nos ha brindado. Este trabajo no solo es el fruto de nuestro esfuerzo, sino también de todas las experiencias y vivencias que hemos tenido a lo largo de este recorrido.

## Resumen

Esta monografía analiza la influencia de las opiniones de figuras públicas en la red social X sobre la gobernanza estatal en el comportamiento humano, utilizando el análisis de sentimientos y la clasificación de textos. A través de una revisión exhaustiva de la literatura, se examinan estudios previos y los resultados de modelos de aprendizaje de máquina aplicados a este tipo de análisis.

El estudio aborda la evolución de las redes sociales, desde ser simples herramientas de comunicación comunitaria hasta convertirse en plataformas de gran impacto en la formación de la opinión pública y la política. Se revisan casos clave, como las elecciones de 2016 en Estados Unidos y la Primavera Árabe, para entender cómo los actores políticos aprovechan estas plataformas para influir en las percepciones sociales y políticas. Además, se exploran los modelos de aprendizaje de máquina que permiten identificar patrones complejos en textos, facilitando la interpretación de las opiniones en tiempo real, así como las limitaciones de su aplicación en la red social X.

**Palabras claves:** Influencia política, Red social X, Análisis de sentimientos, Clasificación de textos, aprendizaje de máquina.

## Abstract

This monograph analyzes the influence of the opinions of public figures in the social network X on state governance on human behavior, using sentiment analysis and text classification. Through a comprehensive literature review, previous studies and the results of machine learning models applied to this type of analysis are examined.

The study addresses the evolution of social networks from being simple community communication tools to high-impact platforms for shaping public opinion and politics. Key cases, such as the 2016 US elections and the Arab Spring, are reviewed to understand how political actors leverage these platforms to influence social and political perceptions. In addition, machine learning models that allow identifying complex patterns in texts, facilitating the interpretation of opinions in real-time, are explored, as well as the limitations of their application in social network X.

**Keywords:** Political influence, social network X, Sentiment analysis, Text, classification, Machine learning.

## Tabla de Contenido

Introducción .....	10
Planteamiento del Problema .....	12
Objetivos.....	17
Objetivo General .....	17
Objetivo Específicos .....	17
Marco Teórico .....	18
Metodología .....	25
Preguntas de Investigación.....	25
Metodología Monografía.....	26
Criterios de Elegibilidad.....	26
Fuentes de Información .....	28
Estrategia de Búsqueda.....	29
Selección de Estudios .....	29
Extracción de Datos.....	29
Evaluación del Riesgo de Sesgo.....	30
Síntesis de Resultados-Análisis Narrativo.....	31
Análisis de Sensibilidad o Grupos.....	31
Resultados.....	33
Grupo 1 .....	34
La Plataforma de la Red Social .....	34
El Branding de la Figura Pública o Política .....	35
Objetivos de la Influencia Política en la Red Social .....	35

Resultados Observados de los Estudios Revisados .....	36
Análisis Crítico .....	39
Grupo 2.....	42
Definición de Objetivos.....	43
Recogida de Datos .....	43
Preprocesamiento de Datos .....	45
Preentrenamiento de los Posts .....	47
Entrenamiento y Optimización .....	49
Evaluación del Modelo.....	52
Implementación y Monitoreo del Modelo .....	54
Grupo 3.....	55
Sentiment Analysis of Real-World Migraine Tweets for Population Research.....	56
Mind Games: A Temporal Sentiment Analysis of the Political Messages of the Internet	
Research Agency on Facebook and Twitter .....	57
Sarcasm Detection Using Machine Learning Algorithms in Twitter: a Systematic Review.	58
Trump vs Biden: Análisis de Sentimientos Elecciones Presidenciales del 2020 .....	60
Conclusiones .....	63
Recomendaciones .....	65
Referencias.....	67

**Lista de Tablas**

<b>Tabla 1</b> <i>Herramientas para Recolección de Datos</i> .....	43
<b>Tabla 2</b> <i>Técnicas de Preprocesamiento</i> .....	45
<b>Tabla 3</b> <i>Estudios e Investigaciones</i> .....	55

## Lista de Figuras

<b>Figura 1</b> <i>Diagrama de Flujo Metodología PRISMA</i> .....	33
<b>Figura 2</b> <i>Arquitectura General</i> .....	42
<b>Figura 3</b> <i>Frecuencia de Algoritmos AMLA</i> .....	60

## Introducción

Este trabajo analiza el impacto en el comportamiento humano, social y político de los usuarios de la red social X, cuando figuras públicas expresan sus opiniones en temas de gobernanza estatal, a través de una revisión de literatura que incluye casos reales, estudios, tendencias y la aplicación de modelos de aprendizaje de máquina para el análisis de sentimientos y clasificación de textos, y que respondan a la pregunta: ¿Cómo influyen las opiniones políticas expresadas por figuras públicas en la red social X sobre la gobernanza, en la comunicación política, la participación ciudadana y la formación de opiniones en un contexto global marcado por la polarización y la desinformación?.

Es importante destacar que Twitter cambió su nombre a X en julio de 2023, como parte de un proceso de rebranding liderado por Elon Musk tras adquirir la compañía. Este cambio no solo incluyó una modificación en el nombre y el logo, sino también una serie de transformaciones en las funcionalidades, objetivos y políticas de la plataforma. Estos ajustes han impactado tanto el servicio como la forma en que los usuarios interactúan con la red social. La actualización de terminología, la incorporación de nuevas características e incluso la permisividad de contenido violento y sexual bajo ciertas condiciones han alterado significativamente el panorama de cómo y con qué propósito se utiliza X.

A diferencia de otras redes sociales como Facebook, Instagram o TikTok, X se enfoca principalmente en contenido textual y conversaciones en tiempo real, además de ser conocida como la red de opinión por su formato de microblogging que permite a los usuarios compartir pensamientos, noticias, y actualizaciones de todo tipo de eventos. X permite el contenido visual pero el texto es su protagonista, y esto abre las puertas para aplicar modelos de aprendizaje de máquina que analicen el texto utilizando procesamiento de lenguaje natural.

Esta monografía sigue la metodología PRISMA, que incluye preguntas de investigación, criterios de elegibilidad, fuentes de información, estrategias de búsqueda, selección de estudios, extracción de datos para realizar una correcta depuración de la literatura revisada y responder a los objetivos planteados.

## Planteamiento del Problema

Las redes sociales han evolucionado desde sus inicios a finales del siglo XX, plataformas como SixDegrees 1997, Friendster 2002 y Myspace 2003, que sentaron las bases para las actuales redes sociales como Facebook, Instagram, X, TikTok, YouTube y LinkedIn. Estas plataformas han crecido y se han diversificado, pasando de ser espacios de conexión y entretenimiento a convertirse en herramientas clave para la comunicación política y social.

Esta transformación ha sido acompañada por un aumento significativo en la adopción de internet a nivel global: dos tercios de la población mundial está conectada, aunque aún existen regiones con grandes brechas digitales, como India y China, donde 680 millones y 336 millones de usuarios respectivamente siguen desconectados. A pesar de esto, la penetración de internet varía considerablemente entre regiones, mientras que EE. UU. 96.8%, en África Central 32.1%, el sudeste asiático 51.5%, y Sudamérica 82.5%.

En 2003, las redes sociales contaban con 150 millones de usuarios activos, cifra que creció exponencialmente a 5.04 billones en 2024, aunque no representa usuarios únicos, ya que una persona puede estar registrada en varias plataformas. En Colombia, el tiempo promedio dedicado a redes sociales es de 3 horas y 46 minutos al día, según Statista y Colprensa (2023), superando el promedio global de 2 horas y 23 minutos. Aproximadamente el 69% de la población colombiana utiliza alguna red social (Arroyabe, s.f.), aunque en países como Corea del Norte, Turkmenistán y Eritrea el acceso está restringido por políticas gubernamentales.

El porcentaje de uso de internet entre los usuarios de 16 a 64 años es de 94.7% en plataformas de chat y mensajes, el 94.3% interactúa en redes sociales y el 80.7% en búsquedas en internet. Para los jóvenes de 16 a 24 años, las redes sociales son la actividad más frecuente (We Are Social, 2024). X, es popular entre usuarios con alto nivel educativo y un interés por la

política y las noticias. Según el Instituto Reuters (Robertson, 2023), el uso de X para mantenerse informado se ha mantenido estable en la última década, siendo más común en Asia y América Latina que en Europa continental.

Es importante destacar que, en julio de 2023, Twitter cambió su nombre a X, tras ser adquirida por Elon Musk en 2022. Esta red fue lanzada en 2006 creada por Jack Dorsey, Biz Stone y Evan Williams, como plataforma de microblogging, X se convirtió rápidamente en una de las plataformas más populares con más de 500 millones de usuarios. Bajo la dirección de Musk, X experimentó cambios significativos, no solo en su nombre y logo, sino también en nuevas funcionalidades como llamadas de video y voz, y el uso de términos como "publicaciones" en lugar de "tuits" y "reposts" en lugar de "retweets". Además, se introdujeron tecnologías de inteligencia artificial y los cambios más notables tras la adquisición incluyen una menor moderación de contenido, promoviendo una "libertad de expresión" que permitió mayor libertad para opiniones y discursos controversiales, como teorías conspirativas, discursos de odio y desinformación que antes podían ser censurados (Silberling, A., Corral, C., y Stringer, A. 2024). También se rehabilitaron cuentas suspendidas, como la de Donald Trump, que había sido bloqueada por violaciones graves de las normas. En 2023 se permitió un mayor rango de contenido explícito, incluyendo imágenes y videos que antes eran prohibidos, este tipo de contenido está sujeto a ciertas condiciones y restricciones, como el etiquetado de material sensible (Es, M. 2024).

El uso de las redes sociales puede tener tanto consecuencias positivas como negativas, y estas dependen en gran medida de cómo y con qué nivel de conciencia las utilice cada persona. Como herramientas de acceso a la información, interacción global en tiempo real y capacidad de respuesta instantánea, las redes sociales ofrecen grandes beneficios. Son especialmente útiles

para mantenernos informados, conectar con otras personas y acceder a una amplia variedad de recursos y contenidos educativos. Sin embargo, cuando su uso no está equilibrado, pueden surgir efectos negativos. Entre ellos se encuentran la desconexión social, el aislamiento, el sedentarismo y la adicción, los cuales afectan tanto nuestra vida social como nuestra salud física y mental. Además, el uso excesivo de redes sociales puede llevar a una comparación social negativa, generando inseguridades y afectando el bienestar emocional. También pueden presentar riesgos en términos de privacidad y seguridad, convirtiéndonos en blancos fáciles para manipulaciones o engaños. Estudios han demostrado que el uso prolongado de las redes sociales está asociado con problemas de salud mental, como ansiedad, depresión y trastornos del sueño, problemas que muchas veces no somos capaces de identificar hasta que se vuelven graves (Kuss & Griffiths, 2017).

La participación política en las redes sociales es vista de manera ambigua por los ciudadanos, según una encuesta del Pew Research Center realizada en 19 economías avanzadas (Wike et al., 2022) en general el 57% de los encuestados considera que las redes sociales han tenido un impacto positivo en la democracia, mientras que el 35% opina lo contrario. Sin embargo, las opiniones varían considerablemente entre países. Un caso destacado es el de Estados Unidos, donde solo el 34% de los adultos cree que las redes sociales han sido beneficiosas para la democracia, y el 64% considera que su impacto ha sido negativo. Otro hallazgo importante de la encuesta es que el 84% de los encuestados en los 19 países cree que el acceso a internet y a las redes sociales ha facilitado la manipulación de la información mediante noticias falsas y rumores. Además, el 65% de los participantes considera que las redes sociales han incrementado la polarización política, y más de cuatro de cada diez opinan que han

empeorado la calidad del debate político, volviendo a las personas menos civilizadas en sus discusiones.

X se ha consolidado como una red social centrada en la opinión, respaldada por estudios como el de Hutchinson (2022). La encuesta realizada a más de 11,500 periodistas en EE. UU., X es la plataforma más utilizada por este grupo, lo que implica que gran parte de la información que se difunde en los medios está, en cierto modo, influenciada por las opiniones expresadas en esta red social. La naturaleza en tiempo real de los posts convierte a X en el espacio ideal para mantenerse actualizado sobre las últimas noticias a medida que ocurren.

Un claro ejemplo de la problemática que surge de la interacción entre políticos, ciudadanos y redes sociales lo encontramos en hechos recientes, como cuando el expresidente de EE. UU., Donald Trump, tomaba decisiones políticas esenciales a través de posts. Asimismo, la influencia de actores externos, como Rusia, en las elecciones presidenciales de 2016 mediante redes sociales pone de manifiesto el poder de estas plataformas. En Colombia, el actual gobierno, al igual que otros en el pasado, ha utilizado X como el medio oficial para comunicar noticias de interés nacional, lo que ha generado divisiones políticas, disputas entre seguidores de distintos bandos y, en algunos casos, incluso escenarios de reconciliación.

El creciente uso de las redes sociales ha transformado estas plataformas en poderosas herramientas políticas, facilitando la comunicación e interacción entre políticos y ciudadanos. Esta interacción influye profundamente en el comportamiento humano, dado que somos susceptibles a las opiniones ajenas y la formación de sesgos. Sin embargo, también recae sobre el ciudadano la responsabilidad de validar la información, conocer a quiénes interactúa y adoptar una postura crítica, respetando las opiniones contrarias. Las redes sociales no son solo espacios de ocio, sino que afectan nuestra capacidad de respuesta mental, física y social a los estímulos,

impulsando la globalización de la comunicación y brindando un sentido de empoderamiento político. En un momento donde muchos se sienten marginados por los sistemas políticos tradicionales, estas plataformas ofrecen un espacio para la participación, permitiendo que los ciudadanos expresen sus voces y se involucren en los temas de interés nacional. Con lo situación encontrada y relatada se ha identificado el siguiente problema:

¿Cómo influyen las opiniones políticas expresadas por figuras públicas en la red social X sobre la gobernanza, en la comunicación política, la participación ciudadana y la formación de opiniones en un contexto global marcado por la polarización y la desinformación?

## **Objetivos**

### **Objetivo General**

Evaluar el impacto de las opiniones de figuras públicas en la red social X sobre la gobernanza estatal en el comportamiento humano, mediante una revisión de la literatura que abarque estudios y análisis de resultados de modelos de aprendizaje de máquina ya desarrollados.

### **Objetivo Específicos**

Identificar el impacto de la red social X en la gobernanza estatal, analizando estudios de casos, referencias bibliográficas, tendencias y patrones emergentes en la plataforma, así como datos históricos relacionados con las opiniones de figuras públicas.

Revisar estudios y trabajos que apliquen modelos de aprendizaje de máquina de análisis de sentimientos y clasificación de textos en la red social X, para entender las metodologías y resultados obtenidos.

Explorar cómo otros proyectos han implementado modelos de análisis de sentimientos y clasificación de textos en la red social X, a partir del conocimiento de las técnicas existentes y su contribución al análisis político y social.

## Marco Teórico

La comprensión de la influencia política en las redes sociales requiere un análisis multidisciplinario que abarca áreas como la comunicación, la ciencia política, la sociología y la informática. En este sentido, nos adentramos en diversas teorías y conceptos que arrojan luz sobre la dinámica de la influencia política en la red social X y el papel de los modelos de aprendizaje de máquina en este contexto.

La evolución de las redes sociales desde finales de los 90 y principios de los 2000 efectivamente revolucionó la recolección y análisis de datos de usuarios, permitiendo a las plataformas estudiar los patrones de comportamiento y optimizar tanto la experiencia del usuario como la publicidad (Fernández, 2024). Con el auge de Facebook, Twitter e Instagram en la década de 2010, el análisis de comportamiento avanzó notablemente gracias al aumento en la cantidad de datos y al perfeccionamiento de herramientas de análisis de datos, como Google Analytics, Apache Hadoop y Tableau, además de algoritmos de aprendizaje de máquina que permiten extraer patrones y predicciones con mayor precisión.

Estos algoritmos de aprendizaje de máquina, que comenzaron a desarrollarse mucho antes, permiten a las plataformas clasificar, predecir y ajustar estrategias de contenido en tiempo real. Fernández (2024) destaca que sus orígenes se remontan a la década de 1950, cuando Arthur Samuel desarrolló un programa de ordenador para jugar a las damas, estableciendo una base para los futuros avances en inteligencia artificial y análisis predictivo. Esta combinación de redes sociales y aprendizaje de máquina ha creado un ecosistema donde se recopilan, analizan y aprovechan los datos de manera continua para impulsar la innovación en experiencias digitales y publicidad segmentada.

La popularidad de las redes sociales impulsó a esta industria a desarrollarse más rápidamente que otros sectores, aprovechando avances en software de análisis de datos, algoritmos de aprendizaje de máquina, sistemas de gestión de contenido (CMS) y mejoras en hardware, como servidores de alta capacidad y dispositivos móviles cada vez más potentes. Este crecimiento también fue respaldado por un recurso humano cada vez mejor preparado académicamente y por una gran cantidad de información disponible, tanto para investigaciones como para el principal motor de negocio detrás de las redes: la publicidad. Esta popularidad convirtió a empresas, gobiernos y personas en actores activos en estas plataformas. Las redes sociales dejaron de ser únicamente medios de comunicación para comunidades o personas de nichos específicos, transformándose en un espacio donde cada actor podía llegar a un amplio público con fines diversos, incluyendo comerciales, sociales o políticos.

Las técnicas de análisis de sentimientos y clasificación de texto tuvieron sus primeras aplicaciones en métodos léxicos simples, que surgieron antes de las técnicas más avanzadas de hoy en día. Como explica Navarro (2023), estos métodos iniciales utilizaban listas predefinidas de palabras con connotaciones positivas o negativas. La idea básica consistía en contar la frecuencia de palabras "positivas" y "negativas" dentro de un texto para estimar su polaridad general. Aunque estos enfoques eran rudimentarios y a menudo imprecisos, representaron un primer paso importante en la cuantificación de sentimientos y opiniones (Navarro, 2023, p. 18).

Posteriormente, estas técnicas evolucionaron gracias a las bases de la lingüística computacional y el procesamiento del lenguaje natural (Jurafsky & Martin, 2019), combinando el poder de las máquinas, el aprendizaje estadístico y su capacidad de procesamiento. Esto permite que los algoritmos aprendan patrones complejos a partir de datos de entrenamiento y apliquen ese conocimiento para clasificar nuevos textos con precisión. Actualmente, literatura

como la de Jurafsky y Martin (2019) que proporciona una visión detallada de los enfoques y técnicas empleadas para extraer información semántica y emocional de textos digitales.

Esta combinación de plataformas de redes sociales y tecnología ha abierto oportunidades para entender, influir, mejorar o alcanzar diversos objetivos en el entorno digital. La aplicación de modelos de aprendizaje automático en redes sociales ha demostrado la eficacia de estos enfoques para comprender y predecir el comportamiento de los usuarios en línea. Estos estudios proporcionan un marco teórico sólido para nuestro análisis de la influencia política en la red social X mediante el uso de modelos de aprendizaje automático, como el análisis de sentimientos con redes neuronales recurrentes (RNN), clasificación de textos con máquinas de vectores de soporte. (SVM) y el modelado de temas con algoritmos como Latent Dirichlet Allocation (LDA)

A nivel empresarial, Criado y Villodre (2018) realizaron un estudio de análisis de sentimientos en X durante el evento de Black Friday en España. En este estudio, se analizaron los posts de 40 empresas con el objetivo de clasificar las percepciones del público sobre estos mensajes como positivos, neutros o negativos. Los resultados mostraron que el 60,2% de los posts fueron clasificados como neutros, el 32,1% como positivos y el 7,7% restante como negativos. Los autores concluyeron que las empresas no aprovechan completamente X para promocionar sus productos durante esta fecha comercial.

Actualmente, las personas están más conectadas y pasan más tiempo navegando en internet que hace algunos años. En el contexto de las redes sociales, esto ha llevado a que las empresas se vean obligadas a identificar nuevos modelos de negocio. Como señalan Moreno, Camacho y Meléndez (2018), en un estudio donde se utilizaron X y herramientas de Python, se evaluaron posts y hashtags mediante modelos de procesamiento de lenguaje natural, con el fin de obtener perspectivas que apoyen la toma de decisiones.

En el ámbito político y gubernamental, encontramos actores que no solo buscan como beneficio tener un canal de comunicación con la población, sino también obtener poder, reconocimiento, formar opiniones políticas, polarizar ideologías o influir en otros. Bossetta (2018) realizó un estudio en el que compara las plataformas Facebook, X, Instagram y Snapchat en relación con la comunicación política y la arquitectura digital durante las elecciones de 2016 en Estados Unidos. Bossetta concluye que la comunicación debe adaptarse a cada plataforma, ya que utilizar el mismo texto, formato y contenido en todas puede debilitar la efectividad de la estrategia de campaña; es preferible personalizar el mensaje según la red social empleada. Este análisis se complementa con el estudio de Sahly, A., Shao, C., y Kwon, K. H. (2019), que examina las estrategias de comunicación de los candidatos Donald Trump y Hillary Clinton en las elecciones de 2016. En X, Trump empleó mensajes con un enfoque en el conflicto y emociones negativas, mientras que Clinton, en X y Facebook, utilizó un tono positivo y un enfoque en el conflicto, adaptando el mensaje a cada plataforma.

Las redes sociales se han vuelto tan fundamentales en el sector político global que muchos gobiernos las utilizan para ejercer influencia en otros territorios, es decir, en países donde no tienen potestad, control ni mandato. Un ejemplo notable es la influencia ejercida por Rusia a través de su Agencia de Investigación de Internet (IRA, por sus siglas en inglés), que difundió mensajes positivos y negativos en plataformas como Facebook y X durante las elecciones de 2016 en Estados Unidos. Esta práctica no se limita a Estados Unidos; se observa en diversas naciones y territorios donde las personas tienen acceso a internet y redes sociales. Según Van Vliet, Törnberg y Uitermark (2020), el uso de redes sociales en la política y la forma en que los políticos las emplean para comunicarse y ganar apoyo varía considerablemente, pues depende de las características específicas de cada país y su población.

En el ámbito de las personas comunes, las redes sociales fueron creadas para facilitar la conexión, reducir distancias, crear comunidades y difundir información, entre otros propósitos que han evolucionado con el tiempo. Un ejemplo claro es el papel fundamental que desempeñó Facebook en 2010 en Egipto, durante el movimiento conocido como la Primavera Árabe. Según Zeynep Meral K., Yurdagül Meral Ü. (2021), las redes sociales fueron esenciales para informar y coordinar a los activistas en relación con la situación y las protestas en Egipto. Este uso de las redes tuvo un impacto tan significativo que contribuyó a la caída del régimen y a la convocatoria hacia un proceso democrático.

Un caso de estudio relevante fue realizado por Lasri, I., Riadsolh, A. y Elbelkacemi, M. (2023), quienes, a través de un análisis de sentimiento en tiempo real de los posts, clasificaron las opiniones de las personas sobre la calidad de las universidades públicas en Marruecos. Este análisis permitió que los datos y su interpretación fueran utilizados por el Ministerio de Educación, Investigación Científica e Innovación para apoyar la toma de decisiones.

Los análisis de información utilizando modelos de aprendizaje de máquina en redes sociales nos han llevado incluso al ámbito médico, como en el caso de la comprensión de la experiencia que un individuo vive al tener una migraña. Según Deng, H. et al. (2020), quienes realizaron un análisis de sentimiento de personas que expresaban vía posts en X de lo que sentían al experimentar migrañas, sus conclusiones mostraron que los tuits de estas personas reflejaban patrones sentimentales extremadamente negativos durante los inicios de la enfermedad.

Todo lo descrito nos lleva a cuestionarnos sobre qué tan influenciables son las personas que utilizan redes sociales, y se destaca una problemática significativa, especialmente en la red social X, anteriormente conocida como Twitter, conocida como la "red de opinión". Las

opiniones de figuras públicas sobre la gobernanza estatal están teniendo un impacto notable en el comportamiento de los usuarios.

Para responder a esta interrogante, se identificará en esta red social la influencia sobre la gobernanza estatal a través de estudios de casos, referencias bibliográficas, tendencias y datos históricos. También se abordarán estudios sobre el análisis de sentimiento y la clasificación de textos utilizando modelos de aprendizaje de máquina en la red social X. Dado que la naturaleza de esta red es permitir la comunicación rápida y precisa a través de texto (280 caracteres), se considera crucial mantener siempre la información actualizada. Finalmente, se explorará cómo otros proyectos han implementado modelos de análisis de sentimiento y clasificación de textos en la red social X, para entender las posibilidades y limitaciones al abordar estos problemas.

A continuación, se presentan aplicaciones de análisis de sentimientos en la red social X, basadas en ejemplos reales y modelos predefinidos que procesan el lenguaje humano:

El trabajo de Devlin, J., et al. (2018), utilizó la técnica BERT en aprendizaje de máquina, que ha revolucionado el campo de la Procesamiento de Lenguaje Natural (PNL). BERT es un poderoso modelo de lenguaje que aprende representaciones contextuales y se aplica a una variedad de tareas, convirtiéndolo en una herramienta valiosa para los investigadores y desarrolladores de PNL. A medida que BERT sigue evolucionando, se espera que tenga un impacto aún mayor en la comprensión y el procesamiento del lenguaje humano.

Otros trabajos como el de Cambria, E. (2017), sobre Affective computing and sentiment analysis, muestra cómo identificar y clasificar los sentimientos expresados en el texto. Esta capacidad permite capturar automáticamente los sentimientos del público en general sobre eventos sociales, movimientos políticos, campañas de marketing y preferencias de productos, un tema que ha generado un interés creciente tanto en la comunidad científica como en el ámbito

empresarial debido a los desafíos y oportunidades que presenta. Adicionalmente Park, S., et al. (2023), examinaron la dinámica temporal de las apelaciones emocionales en los mensajes de la campaña rusa durante las elecciones de 2016. Su estudio analizó las comunicaciones en Facebook y Twitter, evaluando la emoción en el contenido y la orientación de los mensajes, los cuales pudieron haber influido en la percepción de los usuarios.

Por otro lado, Boulogne, S., & Larsson, A. O. (2023), estudiaron cómo los políticos adaptan sus mensajes para diferentes plataformas. En su análisis de las elecciones en Canadá en 2019, descubrieron que los mensajes y contenidos idénticos generaban diferentes niveles de interacción dependiendo de la plataforma, concluyendo que los políticos deben personalizar su contenido y mensaje según la red social en la que se encuentren.

## **Metodología**

Para llevar a cabo este trabajo y cumplir con los objetivos específicos planteados, adoptamos la metodología PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) como marco para la revisión sistemática. Esta metodología nos permitirá identificar, evaluar y sintetizar la evidencia disponible sobre el tema central de nuestra investigación: la influencia política de la red social X, a través del análisis de sentimientos y la clasificación de textos.

La revisión sistemática es una herramienta que se originó principalmente en el ámbito de la salud y las ciencias sociales, debido a su enfoque explícito y predefinido. Inicialmente, se utilizaba principalmente en la medicina, donde los profesionales recurrían a revisiones sistemáticas para mantenerse actualizados en sus especialidades, así como para desarrollar guías de prácticas clínicas. Con el tiempo, otros sectores, como agencias gubernamentales, periódicos, revistas especializadas e investigadores, comenzaron a adoptar este enfoque. La metodología PRISMA, formalmente establecida en 2009, tiene sus raíces en estudios previos, como el de Mulrow en 1987, quien revisó 50 artículos publicados en revistas médicas en 1985 y 1986, descubriendo que ninguno cumplía con los ocho criterios científicos explícitos, como el control de calidad de los estudios. Este análisis fue el germen de lo que más tarde se conocería como el PRISMA Statement (2024).

### **Preguntas de Investigación**

El planteamiento del problema desarrollado en este trabajo se centra en la siguiente cuestión: ¿Cómo influyen las opiniones políticas expresadas por figuras públicas en la red social X sobre la gobernanza, la comunicación política, la participación ciudadana y la formación de opiniones, en un contexto global caracterizado por la polarización y la desinformación? Con el

fin de profundizar en este enfoque y llevar a cabo una investigación basada en la revisión de literatura y el análisis, se plantean las siguientes preguntas de investigación:

- ¿Qué estudios previos han abordado la influencia política en la red social X?
- ¿Qué investigaciones y documentación existen sobre el uso de algoritmos de aprendizaje de máquina para el análisis de sentimientos y la clasificación de textos en la red social X?
- ¿Qué evidencias o proyectos aplicados han utilizado modelos de análisis de sentimiento y clasificación de textos en la red social X?
- ¿Cómo se comparan, comparten y utilizan los resultados de estos proyectos aplicados o casos de estudio?

### **Metodología Monografía**

A continuación, se detalla de manera clara y estructurada cómo se llevará a cabo el proceso de revisión sistemática en esta monografía.

### ***Criterios de Elegibilidad***

Los siguientes criterios se utilizarán para incluir o excluir estudios y publicaciones en la revisión:

Tipo de estudios o publicaciones: La revisión sistemática abarcará una variedad de fuentes, incluyendo artículos de revistas académicas revisadas por pares, tesis, informes de conferencias, artículos de acceso abierto, así como revisiones sistemáticas y metaanálisis, tanto en inglés como en español. Se dará prioridad a aquellos documentos que presenten investigaciones originales sobre el uso de la red social X en el análisis de sentimientos y el discurso político, asegurando que se seleccionen únicamente aquellos con un alto nivel de calidad y rigor metodológico.

**Población:** La población objeto de esta revisión sistemática comprende investigaciones que emplean la red social X como fuente de datos para el análisis de sentimientos y el estudio del discurso político. Esto incluye estudios que analizan interacciones en la plataforma, mensajes políticos y tendencias de la opinión pública. Se focalizará en investigaciones publicadas entre 2016 y 2024, y se buscarán trabajos que ofrezcan una diversidad de enfoques y metodologías, con el objetivo de proporcionar una visión integral sobre el uso de X en el análisis social y político.

**Comparación:** Se podrán comparar tipos de estudio donde se hable de la influencia en otras redes sociales como Facebook, Instagram, Tik Tok entre otras.

**Resultados:** Identificar cómo se mide el impacto de las opiniones y comportamientos generados o influenciados por figuras públicas en la red social X sobre temas de gobernanza estatal, analizando los cambios, patrones o comportamientos en los usuarios de la plataforma. Además, se explorarán las técnicas de análisis de sentimientos y clasificación de textos aplicadas a modelos de aprendizaje de máquina para estudiar estos impactos, con el fin de comprender la efectividad de estas metodologías en la evaluación de la influencia política en los usuarios de la red social X.

**Temas de la publicación:** Las publicaciones incluidas en esta revisión deben abordar los siguientes temas: *La relación de figuras públicas* (políticos, personas famosas, entre otros) con la gobernanza estatal en las redes sociales, con un enfoque particular en la red social X. Se busca explorar cómo estas figuras influyen en el debate político y en la formación de opiniones sobre la gobernanza a través de sus publicaciones y comportamientos en la plataforma. *El impacto de la red social X* en el comportamiento humano, analizando los cambios que genera en los usuarios a nivel físico, mental y social. Esto incluye cómo las interacciones en la plataforma pueden afectar

las percepciones, actitudes y comportamientos de los usuarios en relación con temas políticos y sociales.

El uso de modelos de aprendizaje de máquina, como la clasificación de textos y el análisis de sentimientos, en la red social X, con el objetivo de identificar y comprender las reacciones emocionales y las opiniones de los usuarios ante publicaciones relacionadas con la gobernanza estatal. Las publicaciones deben incluir estudios o investigaciones que utilicen estos modelos para analizar cómo los usuarios perciben y responden a las interacciones de figuras públicas en temas de política.

### ***Fuentes de Información***

Las búsquedas para la revisión sistemática se realizarán en las siguientes bases de datos y sitios especializados en investigación, que ofrecen acceso a publicaciones académicas y científicas relevantes para el tema de estudio:

- Cambridge University Press – Cambridge Journals
- Cornell University
- Frontiers
- Google Books
- Mendeley
- Oxford internet institute
- PLOS ONE
- PubMed
- Research Gate
- Sage Journals
- Textos

### ***Estrategia de Búsqueda***

La búsqueda de estudios relevantes se llevará a cabo mediante una combinación de términos clave y operadores booleanos, con el objetivo de identificar investigaciones pertinentes para el tema de estudio.

- Estudios realizados en los últimos 9 años, es decir desde 2016 a 2024
- “Análisis de sentimiento” and “Twitter” or “clasificación de texto” and “Twitter”
- “Machine learning algorithms” and “twitter” or “Red Social X”
- “Machine learning algorithms” and “twitter” or “Red Social X” and “Sentiment analysis”
- “Sentimental analysis” and “Twitter” or “Red Social X” and “Politics”
- “Social media impact” and “politics”
- “Twitter” or “Red Social X” and “politics”
- “Natural Language Processing Applications” and “twitter” or “Red Social X”

### ***Selección de Estudios***

Dos revisores evaluarán de forma independiente los títulos y resúmenes de los estudios obtenidos, aplicando los criterios de inclusión establecidos. Los estudios que cumplan con dichos criterios serán seleccionados para una revisión completa. En caso de discrepancias entre los revisores, se recurrirá a un tercer revisor para llegar a un consenso y asegurar la consistencia en la selección de los estudios.

### ***Extracción de Datos***

Se llevará a cabo una revisión detallada de los estudios y publicaciones seleccionadas, con el objetivo de extraer la siguiente información clave:

1. Datos de Identificación

2. Autoría
3. Método
4. Muestra
5. Planteamiento
6. Principales hallazgos
7. Conclusiones

Esta información será organizada y analizada para facilitar la síntesis de los resultados relevantes para la investigación.

### ***Evaluación del Riesgo de Sesgo***

Para asegurar la calidad metodológica de los estudios seleccionados, se realizará una evaluación exhaustiva conforme a los siguientes criterios:

Revisión completa de todos los estudios: Se revisará el 100% de los estudios seleccionados, garantizando que cumplan como mínimo con los requisitos de elegibilidad establecidos.

Análisis de estudios sobre modelos de aprendizaje de máquina: En los estudios que utilicen modelos de aprendizaje de máquina para el análisis de sentimientos o la clasificación de textos, se verificará que el procedimiento metodológico esté claramente definido. Esto incluirá la relación de la muestra utilizada, la descripción de cómo se aplicó el modelo en la red social X, y una conclusión clara sobre los resultados obtenidos.

Estudios sobre la influencia de figuras públicas en la gobernanza estatal: Para los estudios que analicen la influencia de figuras públicas en la gobernanza estatal a través de la red social X, se evaluará que se proporcione una medición clara y una descripción detallada de la técnica

utilizada. Además, se revisará cómo se identificaron los comportamientos o influencias de las figuras públicas y la conclusión a la que llegaron.

Transparencia y crítica en la comunicación de los estudios: Se evaluará que los estudios seleccionados presenten una comunicación clara, transparente y crítica respecto a la metodología utilizada y los hallazgos obtenidos, asegurando que se haya reflexionado sobre las limitaciones y fortalezas de cada investigación.

### ***Síntesis de Resultados-Análisis Narrativo***

Cualitativo: En los casos en que los estudios incluyan datos cualitativos o presenten una gran heterogeneidad en los enfoques, métodos o contextos, se recurrirá a un análisis narrativo. Este enfoque es particularmente útil para explorar los hallazgos de manera más profunda y contextualizada, permitiendo una discusión detallada sobre las implicaciones de los resultados.

Flexibilidad: El análisis narrativo ofrece la flexibilidad necesaria para identificar y explorar temas, patrones y relaciones que no siempre pueden ser cuantificados. Es ideal para resumir y discutir tanto las implicaciones teóricas como prácticas de los resultados obtenidos en los estudios.

Interpretación Subjetiva: Aunque el análisis narrativo no tiene la misma rigurosidad que un metaanálisis, proporciona valiosas perspectivas basadas en la experiencia y juicio del investigador. Permite una interpretación más rica de los datos, considerando las diversas dimensiones del fenómeno estudiado.

### ***Análisis de Sensibilidad o Grupos***

Dependiendo de los resultados obtenidos, se podrán definir los siguientes grupos para un análisis más específico:

Estudios sobre la influencia en la gobernanza estatal en la red social X: Este grupo incluirá aquellos estudios que se centren en cómo las figuras públicas y otros actores influyen en la gobernanza estatal a través de sus interacciones en la red social X.

Estudios que utilizan modelos de análisis de sentimientos y clasificación de texto en la red social X: Este grupo incluirá investigaciones que emplean modelos de aprendizaje de máquina en la red social X, ya sea en su totalidad (desde la captura de datos hasta el análisis de resultados) o en etapas específicas del proceso, tales como la recolección de datos, el almacenamiento, el desarrollo del modelo, su implementación y el análisis de los resultados obtenidos.

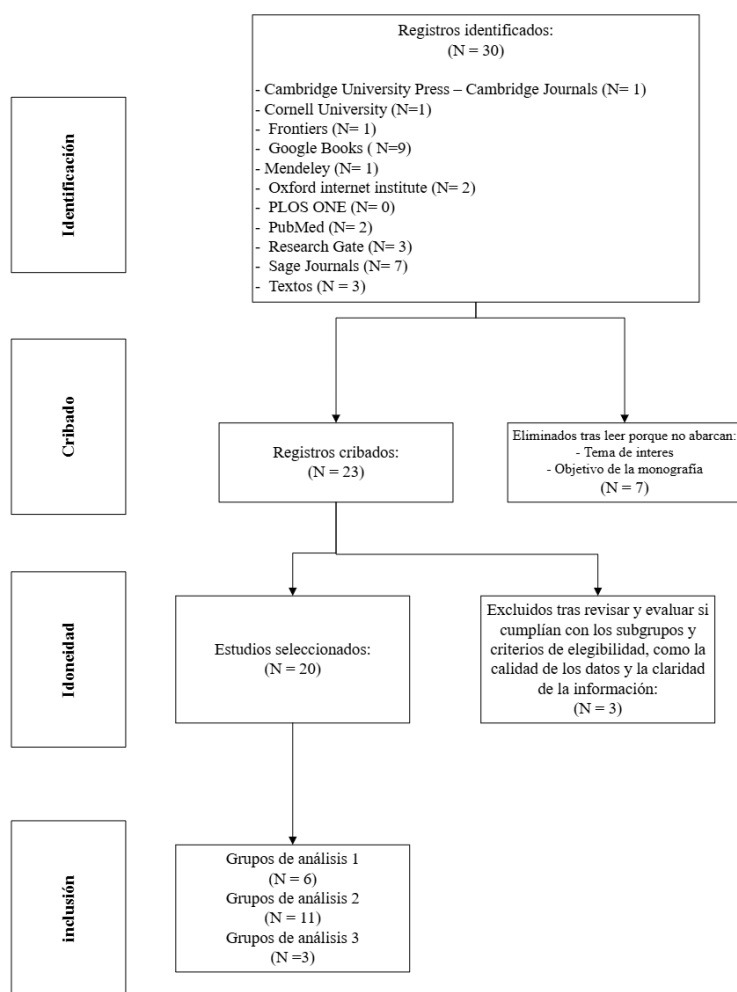
Análisis y resultados de la implementación de modelos de aprendizaje de máquina en la Red Social X: Este grupo recopila los estudios y proyectos aplicados utilizando modelos de aprendizaje de máquina para clasificar texto y análisis de sentimiento en la red social X, a través de una revisión detallada, se busca proporcionar una comprensión profunda de los impactos que estos modelos generan en la interacción de los usuarios y en la dinámica social y política dentro de la red.

## Resultados

Se analizaron un total de 30 investigaciones, artículos, libros y otros documentos académicos que cumplían con los criterios definidos en la estrategia de búsqueda implementada en diversas fuentes de información. Después de aplicar la metodología de selección, se obtuvieron los siguientes resultados, los cuales se presentarán en tres grupos de análisis:

**Figura 1**

### Diagrama de Flujo Metodología PRISMA



*Nota.* Los grupos de análisis 1, 2 y 3 corresponden a los descritos en la sección de análisis de sensibilidad detallada en la metodología.

## **Grupo 1**

Estudios sobre la Influencia de la Red Social X en la Gobernanza Estatal, enfocados en analizar cómo esta red impacta en la gestión del Estado. Se incluirán aquellos trabajos que aborden cómo las figuras públicas y otros actores clave ejercen su influencia sobre la gobernanza a través de sus interacciones en dicha plataforma, destacando el papel de la red social como un espacio de interacción política y social.

Los hallazgos obtenidos a partir de los artículos revisados sugieren que la influencia de actores políticos o figuras públicas en la gobernanza estatal a través de la red social X depende de varios factores clave, que se resumen a continuación:

### ***La Plataforma de la Red Social***

La influencia varía según las características propias de cada plataforma. En el caso de X, se destacan los siguientes aspectos:

- El tipo de networking que se puede generar, influenciado por funcionalidades como "me gusta", reposts, etiquetas, hashtags, hilos, entre otros.
- La creación de diferentes temáticas, como los debates sobre gobernanza estatal, donde los usuarios utilizan la plataforma para participar políticamente, expresar opiniones, intercambiar información y formar grupos según intereses comunes.
- El proceso de cuantificación y cualificación de actividades, características e incluso la segmentación de usuarios, conocido como datificación, que la plataforma pone a disposición de los actores políticos como herramienta para análisis y toma de decisiones.

### ***El Branding de la Figura Pública o Política***

El impacto de una figura pública o política está estrechamente ligado a su imagen, marca personal y al mensaje y tono con los que se comunica. Estos elementos influyen en varios aspectos:

- **Visibilidad:** Mantener su nombre e imagen presentes en la mente de los usuarios, asegurando su constante reconocimiento.
- **Atracción:** Atraer nuevos usuarios a la plataforma, ampliando su audiencia y su influencia.
- **Lealtad:** Fomentar la fidelidad de los usuarios hacia su marca personal, cultivando una base de seguidores comprometidos.
- **Relación e interacción:** La plataforma X facilita una conexión directa entre el político y el usuario, promoviendo la confianza, la autenticidad y una sensación de cercanía, lo que mejora la interacción y la intimidad digital.
- **Mensaje y tono:** El mensaje y el tono con el que se comunica la figura pública son determinantes para generar respuestas distintas en los usuarios. Por ejemplo, los mensajes negativos o los ataques hacia adversarios suelen generar un mayor volumen de comentarios e interacciones.

### ***Objetivos de la Influencia Política en la Red Social***

Los objetivos de la influencia política se refieren a los fines que busca alcanzar una figura pública al expresar su opinión en la red. Estos objetivos incluyen los siguientes aspectos:

- Promover una ideología política.
- Influir en la campaña electoral, buscando movilizar a los votantes con la elección de un candidato.

- Incrementar la cuota de fondos y aumentar las donaciones monetarias destinadas a financiar la campaña política.
- Fomentar la participación política, invitando a los usuarios a involucrarse activamente en la política, promoviendo la acción política mediante la interacción en la plataforma.

### ***Resultados Observados de los Estudios Revisados***

Este estudio explora la dinámica de compartir peticiones gubernamentales en línea a través de X, un tema poco investigado hasta ahora. Analizando más de un millón de posts vinculados a peticiones durante 20 meses, se examinan las redes de peticiones y sus simpatizantes, revelando patrones sociales implícitos, este artículo de Cihon, P., et al. (2021) realizado a usuarios en el Reino Unido tiene como resultados:

- Los usuarios de X no se limitan a compartir peticiones sobre un solo tema ni aquellas que son populares. Entre más de 240.000 usuarios, se identificaron grupos de apoyo latentes, donde los usuarios más influyentes son, en su mayoría, personas políticamente activas, pero no extremas.
- Estos grupos que emergen en la plataforma suelen ser fluidos, ad hoc y temporales, formándose a partir de pequeñas acciones de participación de los usuarios, lo que desafía los modelos tradicionales de organización política basados en la identidad o los intereses compartidos.
- Los resultados sugieren que la participación de los usuarios en la red genera una diversidad de opiniones, intereses y grupos dentro del sistema político y social tan amplia y desorganizada que dificulta la coordinación y estructuración eficiente de las interacciones entre estos actores, aunque, al mismo tiempo, favorece una mayor espontaneidad.

- X demuestra tener el potencial para facilitar la creación y coordinación de nuevos grupos de interés en línea, favoreciendo la participación política y la movilización social.

La participación de Donald Trump como candidato presidencial en las elecciones de 2016 y 2024, ambas resultando en victorias contundentes, pone de manifiesto el impacto positivo de sus estrategias políticas en las redes sociales. Un estudio realizado por Sahly, A., Shao, C., y Kwon, K. H. (2019) analiza las diferencias en las estrategias de comunicación digital entre los dos principales candidatos en las elecciones de 2016, Donald Trump y Hillary Clinton, centrándose en el uso de Facebook y la red social X, para lograr el compromiso de sus audiencias. Los resultados principales del estudio fueron los siguientes:

- Los candidatos emplearon enfoques significativamente distintos en cada plataforma. Mientras que los efectos de ciertos tipos de mensajes (como los relacionados con el conflicto, la moralidad y las emociones) sobre la participación de la audiencia fueron consistentes en X, en Facebook los efectos fueron inconsistentes.

- Trump utilizó una mayor variedad de mensajes, con un énfasis particular en los tonos emocionales negativos y el conflicto en X. En Facebook, por el contrario, se centró en mensajes más positivos. Clinton, en cambio, adoptó una estrategia basada principalmente en el conflicto y en mensajes emocionales positivos en ambas plataformas.

- El estudio encontró que los marcos de conflicto y moralidad estaban fuertemente asociados con el comportamiento de reposts (compartir publicaciones) entre los seguidores de ambos candidatos en X. Por otro lado, los mensajes emocionales tendían a generar una mayor interacción en forma de "me gusta" y favoritismo. En Facebook, sin embargo, no se observó un patrón claro ni consistente en los comportamientos de participación entre los seguidores de los dos candidatos.

- Este estudio subraya el papel cambiante de las redes sociales en las campañas políticas y cómo los candidatos adaptan sus mensajes según la plataforma y la audiencia. Los resultados también sugieren que los mensajes populistas y agresivos de Trump en las redes sociales podrían tener implicaciones duraderas en el discurso político y las normas democráticas.

El artículo de Fatema, S., Yanbin, L., Fugui, D.(2022), analiza la influencia de los medios sociales en la relación entre políticos y ciudadanos, con un enfoque particular en el efecto moderador de los eslóganes políticos de cuatro partidos en Pakistán. A partir de los datos obtenidos de diversos usuarios de redes sociales, el estudio concluye lo siguiente:

- El uso de las redes sociales tiene un impacto positivo y significativo tanto en la relación general entre políticos y ciudadanos como en la interacción entre la política y las redes sociales.
- Los eslóganes de los partidos políticos juegan un papel importante en la creación de un vínculo con los ciudadanos, incentivando su participación política.
- Se acepta la Hipótesis 1, que establece que los medios sociales influyen en la relación entre políticos y ciudadanos.
- De los cuatro eslóganes políticos estudiados, tres demostraron tener una influencia positiva en dicha relación.
- Tres de estos eslóganes también fueron efectivos para influir en el comportamiento de voto de los ciudadanos, favoreciendo a los partidos correspondientes.
- Los partidos políticos y los políticos que están más activos y organizados en las redes sociales tienen mayores probabilidades de llegar a los ciudadanos de manera eficaz.

El estudio Boulianne, S. y Larsson, A. O. (2023). investiga la participación en las redes sociales de los líderes políticos canadienses durante la campaña electoral federal de 2019,

centrándose específicamente en Facebook, X e Instagram. Se analizaron las publicaciones realizadas por los líderes de los tres principales partidos políticos se encontró que:

- Cuando los mensajes contenían ataques hacia otros candidatos, en su mayoría mensajes negativos, tenían una mayor atención e interacción, ya que aumentaban los comentarios y comparticiones.
- El tono del mensaje también influye en el usuario, por ejemplo, los posts mensajes negativos generaban más “me gusta” mientras que los mensajes con ataques generaban más comentarios y comparticiones.
- La interacción puede ayudar a reducir la distancia psicológica entre los políticos y los votantes, haciendo que estos se sientan más comprometidos personalmente con la campaña.
- Las publicaciones interactivas (por ejemplo, etiquetando o mencionando a los usuarios) promueven una sensación de autenticidad e intimidad, que también puede aumentar la participación, pero puede ser más difícil de ejecutar con eficacia.
- Los posts que hacen referencia a políticos internacionales de alto perfil (como Alexandria Ocasio-Cortez o Barack Obama) tendieron a recibir un alto compromiso, lo que sugiere que los usuarios de X son receptivos a figuras políticas globales y debates internacionales.

### ***Análisis Crítico***

En la actualidad, la participación política a través de las redes sociales se ha convertido en un componente esencial para cualquier actor político. Independientemente de sus objetivos, la mayoría de los políticos buscan influir en la opinión pública, sembrando ideas y dejando que la audiencia decida por sí misma cómo interpretarlas. Es aquí donde el mensaje y el tono con los que un político se presenta en estas plataformas cobran una relevancia especial. La manera en

que un político se comunica, y el reconocimiento de su marca personal o ideológica frente a la audiencia, redefine las reglas del juego en la política moderna. En este contexto, las redes sociales no solo actúan como un canal de comunicación, sino que también estructuran y modulan la interacción entre los actores políticos y la sociedad.

Es innegable que los seres humanos somos susceptibles a la influencia externa, y el poder que las redes sociales tienen sobre la creación de opinión es un tema clave en el análisis político contemporáneo. En este sentido, surge una interrogante crucial: ¿están los ciudadanos realmente preparados para generar opiniones objetivas, o caen fácilmente en la trampa de aceptar ideas y dogmas impuestos sin cuestionarlos? Aunque todos tenemos un margen de error, la capacidad de ser influenciados puede poner en riesgo la calidad democrática. Por esta razón, las plataformas sociales deberían jugar un rol activo en proteger la vulnerabilidad humana ante la desinformación, los discursos de odio y otros contenidos tóxicos. Sin embargo, lograr esto es complejo, como se observa en la plataforma X, donde la moderación de contenido es prácticamente nula. En este entorno, los usuarios tienen acceso a una gran cantidad de información, pero la responsabilidad recae en ellos para decidir qué es cierto o falso. Algunos defienden este enfoque de libertad total argumentando que permite a los usuarios formarse su propia opinión sin restricciones externas.

Sin embargo, las redes sociales como X no solo facilitan la interacción política, sino que también estructuran esa interacción. X, en particular, ofrece un espacio donde los usuarios pueden expresar opiniones políticas de manera rápida, clara y directa. Pero esta interacción no está limitada solo a los políticos y ciudadanos; hay un tercer actor fundamental en este proceso: la propia red social. El papel que juegan estas plataformas en la creación de discursos políticos y en la mediación entre ciudadanos y políticos es una cuestión que aún se encuentra

insuficientemente explorada. La manera en que las plataformas gestionan la visibilidad de ciertos contenidos, las características algorítmicas que priorizan ciertos mensajes, y la segmentación de usuarios a través de datos demográficos, intereses y comportamientos, influyen profundamente en la forma en que se desarrollan las conversaciones políticas.

Un ejemplo claro de cómo las plataformas pueden influir en la política es la campaña presidencial de Barack Obama en 2012. Utilizando los datos recogidos de los usuarios en Facebook, la campaña de Obama fue capaz de segmentar a votantes susceptibles de ser persuadidos, animando a sus simpatizantes a enviar mensajes personalizados a sus amigos, lo que les permitió maximizar el impacto de su mensaje político. Este uso de la "datificación" de los usuarios muestra cómo las redes sociales no solo sirven como canales de comunicación, sino que también son herramientas poderosas para modelar el comportamiento electoral.

En conclusión, las redes sociales han transformado la participación política al ofrecer nuevas formas de interacción, pero también han generado nuevos desafíos relacionados con la desinformación y la manipulación de opiniones. La influencia de las plataformas sociales va más allá de ser un simple medio de comunicación; estas plataformas modelan las conversaciones políticas, segmentan a los usuarios y configuran los mensajes que los ciudadanos reciben. Es crucial investigar y entender cómo estos factores afectan la percepción política de los usuarios y cómo las plataformas pueden actuar para mitigar los riesgos asociados con la desinformación. La responsabilidad de las redes sociales es, por tanto, un tema de gran relevancia para la política contemporánea, pues, aunque la libertad de expresión es esencial, la protección frente a la manipulación es igualmente importante para asegurar una democracia saludable y fundamentada.

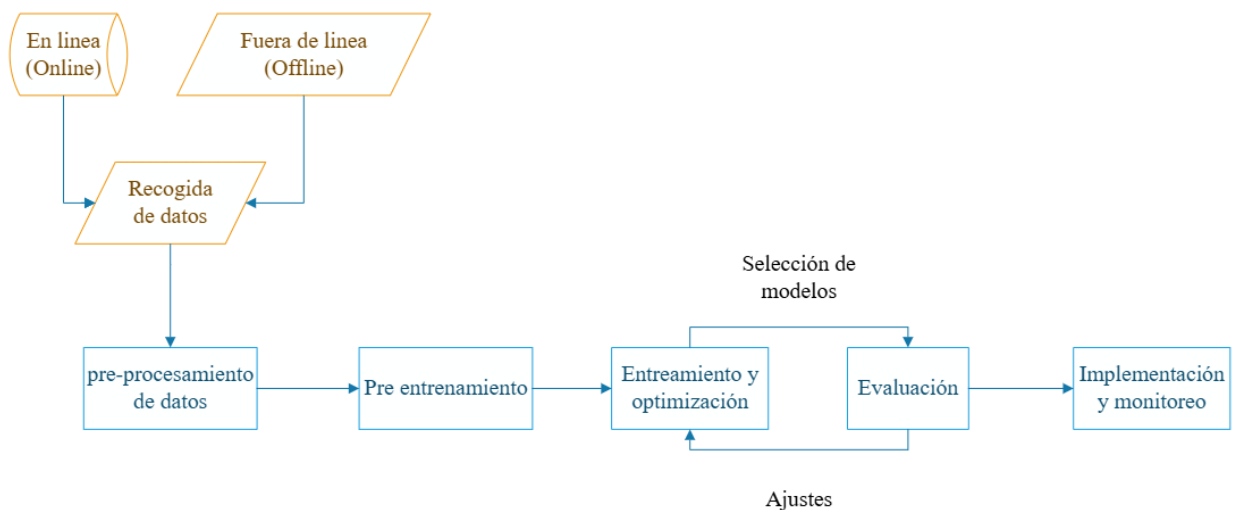
## Grupo 2

Estudios de Análisis de Sentimientos y Clasificación de Texto en la Red Social X, este grupo de análisis agrupa los estudios que aplican estos modelos en la red social X. En este se incluyeron investigaciones que emplean técnicas de aprendizaje de máquina, ya sea en su totalidad, abarcando todo el proceso (desde la captura de datos hasta el análisis de resultados), o en etapas específicas, como la recolección y almacenamiento de datos, el desarrollo del modelo, su implementación y la interpretación de los resultados obtenidos.

A continuación, se presenta una arquitectura general para implementar un modelo de clasificación o análisis de sentimiento en la red social X, ver ilustración 2. Es importante destacar que cada estudio puede requerir una arquitectura personalizada para abordar necesidades específicas:

### Figura 2

#### *Arquitectura General*



*Nota.* Arquitectura general para implementar un modelo de análisis de sentimiento y clasificación de texto.

La arquitectura propuesta de la figura 2 aborda los siguientes pasos:

### ***Definición de Objetivos***

Definir qué tipo de análisis se desea identificar, positivo, negativo, neutral y que categorías son relevantes para lograr el objetivo, adicional validar si lo que se identifica va a ser un input para ser utilizados como información en otros modelos.

### ***Recogida de Datos***

Según la disponibilidad, los datos pueden ser recolectados de dos maneras: en línea (es decir, en tiempo real, a medida que se genera la información) o fuera de línea (cuando los datos ya han sido recopilados y almacenados sin haber sido procesados en tiempo real). Tal como se menciona en el artículo de Chen, K., Duan, Z., y Yang, S. (2022), entre las herramientas disponibles en el mercado se incluyen:

**Tabla 1**

### ***Herramientas para Recolección de Datos***

Herramienta	Descripción	Precio	Consideraciones Adicionales
API de X	Permite acceder a posts, perfiles, mensajes directos, listas, tendencias, y lugares. Ideal para análisis y clasificación de texto.	Gratuita (limitada) Básica (desde USD\$ 175/mes) Pro (desde USD \$4.500/mes) Premium (Sin información)	X ofrece cuatro versiones: gratuita, Básica, Pro y Enterprise, cada una con diferentes restricciones y funcionalidades. La versión gratuita tiene limitaciones en el número de solicitudes y tipos de datos, mientras que las versiones premium ofrecen más acceso y características. Más información en <a href="https://developer.x.com/en/docs">https://developer.x.com/en/docs</a> .
Third-party platforms	Son herramientas externas que facilitan la interacción entre usuarios y negocios como Hootsuite, Sprout Social o Brandwatch que ofrecen acceso a datos de X y análisis integrados.	Varía según la plataforma, desde \$15/mes hasta varios cientos.	Proporcionan interfaces amigables y análisis avanzados, pero pueden no ofrecer acceso completo a todos los datos.

Posts IDs compartidos (metadata)	Captura de IDs de posts compartidos por otros usuarios para análisis posterior.	Gratis	Necesita acceso a los posts originales para obtener contenido completo; limitado a lo que se comparte públicamente.
Scraping (web scraping)	Técnica para extraer datos directamente de la interfaz web de X. Puede ser útil para capturar contenido no disponible a través de la API.	Gratis (pero puede ser ilegal o violar términos de servicio)	Requiere habilidades técnicas y puede ser riesgoso en términos legales y éticos. No recomendado si se puede evitar.
Data dumps o data sets públicos	Conjuntos de datos previamente recopilados y compartidos por investigadores o empresas. Pueden contener posts históricos.	Generalmente gratis	Útil para análisis históricos, pero puede no tener datos recientes o específicos.

---

*Nota.* Herramientas para recolección de datos en análisis de Sentimiento y clasificación de texto.

Para tener en cuenta:

- Entender el beneficio y alcance de cada herramienta para capturar datos puede ser crucial para poder procesar la información que se captura.
- La necesidad de contar con personal técnico para capturar, limpiar e implementar la información de X puede ser una limitante en este tipo de investigaciones, cabe aclarar que las herramientas de terceros tratan de ayudar al investigador creando sitios más fáciles e intuitos de manejar, pero suelen ser más costosos.
- Asegurarse de cumplir con las políticas de uso de X y las leyes sobre privacidad y protección de datos al usar cualquier método para capturar información.
- Cada método tiene sus propias limitaciones en cuanto a la cantidad y tipo de datos que se pueden obtener, así como la frecuencia con la que se pueden realizar las solicitudes.
- Cuando se tienen modelos online, no se tiene el control sobre la información que captura y tampoco el costo de obtenerla.

- Las herramientas de captura de datos no garantizan el 100% de los posts sobre la categoría configurada y la mayoría tampoco garantiza que se capture de manera aleatoria, esto es difícil de identificar y si puede generar inconvenientes al entrenar los modelos.
- La captura de datos presenta un gran reto y es el diferenciar cuando un post lo realiza un humano o un Bot, esto conlleva a que la investigación pierda confianza.

### ***Preprocesamiento de Datos***

Es una de las etapas más críticas en cualquier proyecto de aprendizaje automático incluido el análisis de sentimientos es transformar y limpiar los datos crudos para hacerlos más útiles y consistentes, adecuándolos a los algoritmos de modelado. Un preprocesamiento adecuado es fundamental, ya que, sin él, el modelo podría generar resultados incorrectos o ineficaces. Con base en los artículos de Sarsam, S. et al (2020), Karaman Lheureux, Yasemin (2023) y Lasri, I., Riadsolh, A., y Elbelkacemi, M. (2023) se muestra un consolidado de las técnicas más comunes en este tipo de modelos de análisis y clasificación de texto:

**Tabla 2**

#### *Técnicas de Preprocesamiento*

Preprocesamiento	Descripción	Notas Adicionales
Concatenación de Cadenas	Combina tokens nuevamente en una sola cadena después del procesamiento, si es necesario para un análisis posterior.	Se hace después de la tokenización, lematización y otros pasos al preparar el texto final para el análisis.
Conversión a Minúsculas	Convierte todo el texto a minúsculas para asegurar uniformidad y prevenir discrepancias sensibles a mayúsculas.	Ayuda a reducir la dimensionalidad de los datos de texto.
Corrección de Palabras Negativas	Ajusta palabras afectadas por la negación (por ejemplo, cambiar "no feliz" a "infeliz") para mejorar la detección del sentimiento.	Ayuda a mejorar el análisis de sentimiento entendiendo mejor las negaciones.
Corrección ortográfica	Identifica y corrige palabras mal escritas en el texto, mejorando la calidad general para el análisis.	Útil cuando se trabaja con texto informal, como los posts en X.

Eliminación de Caracteres Especiales	Elimina caracteres no alfanuméricos (@, #, %, &, etc.) que no aportan valor al análisis.	Generalmente se realiza con expresiones regulares.
Eliminación de Caracteres Irrelevantes	Elimina caracteres que no contribuyen al significado del texto, como puntuaciones y símbolos.	Generalmente se hace usando expresiones regulares (regex).
Eliminación de Duplicados	Elimina entradas duplicadas en el conjunto de datos para asegurar instancias únicas para el análisis.	Asegura que las palabras repetidas no distorsionen el análisis.
Eliminación de Hashtags	Remueve hashtags del texto, que pueden no ser necesarios para la extracción de sentimientos.	
Eliminación de Menciones	Borra menciones de usuarios (por ejemplo, "@usuario") que no contribuyen al sentimiento que se está analizando.	Especialmente importante en datos de redes sociales como X, Instagram, Facebook.
Eliminación de Palabras Vacías	Remueve palabras comunes (por ejemplo, "y", "el", "es") que no añaden un significado significativo al análisis.	Ayuda a reducir el espacio de características y enfoca el análisis en palabras más significativas.
Eliminación de Reposts	Excluye indicadores de reposts (por ejemplo, "RT") y duplicados causados por reposts.	Esto es especialmente relevante para datos de X.
Eliminación de URLs	Elimina direcciones web del texto, que generalmente son irrelevantes para el análisis de sentimientos.	Normalmente se hace mediante la detección de patrones de URL con expresiones regulares.
Etiquetado de partes del discurso (POS)	Asigna partes del discurso (por ejemplo, sustantivo, verbo, adjetivo) a cada token, lo que puede ayudar a comprender el contexto.	Útil para tareas como el análisis sintáctico y la selección de características para procesamiento posterior.
Lematización	Similar a la reducción, pero reduce las palabras a su forma base según definiciones del diccionario (por ejemplo, "mejor" a "bueno").	Más precisa que el stemming, ya que utiliza un diccionario y el significado de las palabras.
Manejo de Emojis y Emoticonos	Convierte emojis y emoticonos en representaciones textuales o los elimina según su relevancia.	Puede ser importante para el análisis de sentimiento, ya que los emojis suelen expresar emociones.
Reducción (Stemming)	Reduce las palabras a su forma base o raíz, a menudo eliminando sufijos (por ejemplo, "corriendo" a "correr").	
Tokenización	Divide el texto en palabras o tokens individuales, facilitando su análisis.	Puede hacerse a nivel de palabra (tokens de palabra) o a nivel de frase (tokens de frase).

*Nota.* Técnicas de preprocesamiento en modelos de análisis de sentimiento y clasificación de texto.

### *Preentrenamiento de los Posts*

El preentrenamiento en procesamiento de lenguaje natural (PLN) se refiere a técnicas que permiten representar el texto en una forma numérica (vectores) que los modelos de aprendizaje automático pueden procesar. Esto permite a los modelos captar relaciones semánticas y sintácticas complejas entre palabras y frases, mejorando su capacidad para tareas de clasificación Sanhueza Arancibia, R. I. (2013) y Lasri, I., Riadsolh, A., y Elbelkacemi, M. (2023). Algunas de estas técnicas de preentrenamiento en clasificación de texto son:

- **Word2Vec:** es una técnica de preentrenamiento desarrollada por Google que representa palabras como vectores densos en un espacio de alta dimensionalidad. Existen dos enfoques principales en Word2Vec: Skip-gram que predice las palabras contextuales a partir de una palabra central y CBOW (Continuous Bag of Words) que predice una palabra central a partir de su contexto.
- **GloVe (Global Vectors for Word Representation):** Esta técnica se basa en la matriz de co-ocurrencia de palabras en el corpus. En lugar de utilizar una red neuronal como Word2Vec, GloVe factoriza esta matriz para obtener representaciones vectoriales. Los vectores resultantes son también capaces de capturar relaciones semánticas entre palabras.
- **BERT (Bidirectional Encoder Representations from Transformers):** desarrollado por Google, es un modelo basado en la arquitectura transformer que se pre-entrena de manera bidireccional, es decir, tiene en cuenta el contexto tanto de la izquierda como de la derecha de una palabra en una secuencia. Esto permite que BERT capture relaciones de contexto más complejas y precisas, lo que lo hace altamente efectivo para tareas como clasificación de texto, análisis de sentimientos, y respuestas a preguntas.

- GPT (Generative Pre-trained Transformer): GPT es otro modelo basado en transformer, pero se entrena de manera unidireccional (de izquierda a derecha). Aunque es más conocido por su uso en generación de texto, también puede ser adaptado para tareas de clasificación, análisis de sentimientos, etc. GPT-3, por ejemplo, es capaz de realizar una amplia variedad de tareas sin necesidad de un entrenamiento adicional específico gracias a su preentrenamiento masivo.

Técnicas con librerías de Python:

- TextBlob: Es una librería de Python que proporciona herramientas simples para realizar tareas de procesamiento de lenguaje natural (PLN) como análisis de sentimiento, traducción y corrección gramatical. Aunque TextBlob no es una técnica de preentrenamiento por sí sola, utiliza modelos pre-entrenados como los de NLTK y pattern para facilitar tareas de procesamiento de texto.

- spaCy: Una librería de Python avanzada que ofrece modelos pre-entrenados para una variedad de tareas de PLN, como análisis sintáctico, extracción de entidades, y clasificación de texto. spaCy es conocida por su velocidad y eficiencia.

- Transformers, Wolf, T. et al (2020): Esta librería permite acceder a una variedad de modelos pre-entrenados basados en Transformers como BERT, GPT, RoBERTa, T5, entre otros. Estos modelos pueden ser fácilmente adaptados para tareas específicas de PLN, como clasificación de texto, generación de texto, análisis de sentimientos, y mucho más.

- Para tener en cuenta: Algunas técnicas como BERT y GPT son muy poderosas, también suelen ser muy costosas en términos de recursos computacionales, especialmente en tareas de gran escala. Las técnicas de preentrenamiento pueden reflejar sesgos en la construcción de los datos que pueden influir en las tareas de clasificación.

### ***Entrenamiento y Optimización***

En esta etapa se identifica las características semánticas más relevantes de los posts (variables predictoras) que permitan al modelo realizar predicciones precisas. Para que un modelo funcione de manera óptima y logre un mejor rendimiento al clasificar el post como positivo, negativo o neutral, es necesario convertir la información cualitativa en cuantitativa, es decir, transformar los datos textuales en características numéricas. Esto facilita a los analistas obtener una mayor precisión y comprensión sobre el sentimiento público en diversos dominios, como se señala en el artículo de Lasri, I., Riadsolh, A., y Elbelkacemi, M. (2023). En este proceso se utilizan técnicas como:

- Count Vectorizer y TF-IDF (Term Frequency-Inverse Document Frequency), que permiten vectorizar el texto y convertir las representaciones semánticas en características numéricas.
- Ajuste de hiperparámetros, con el fin de optimizar el rendimiento del modelo, mejorar su capacidad de generalización y aumentar la eficiencia del entrenamiento, lo que se traduce en mejores resultados.
- Validación cruzada, que busca aumentar la confianza en el modelo, reducir el sobreajuste, mejorar la eficiencia cuando los datos son limitados, ajustar los hiperparámetros y comparar diferentes modelos para seleccionar el más adecuado.

Los modelos utilizados para el análisis de sentimiento y clasificación de texto se agrupan en las categorías: tradicionales adaptados, tradicionales personalizados y avanzados. Todos estos modelos son supervisados y se utilizan para clasificación, como se menciona en los estudios de Sarsam, S. et al. (2020), Karaman Lheureux, Yasemin (2023) y Devlin, J., et al. (2018).

Modelos tradicionales adaptados:

- Support Vector machine (SVM): Facilita la clasificación de sentimientos en posts como positivo, negativo y neutro.
- Random Forest: clasifica los textos en positivo, negativo y neutro.
- Naive Bayes: este es usado para clasificar la polaridad de los posts es decir positivo o negativo.
- Logistic regression: Es útil para realizar una predicción binaria, es decir, positivo o negativo.

Modelos tradicionales personalizados: Existen diversos modelos utilizados en el análisis de sentimientos y clasificación de texto, los cuales se agrupan en modelos supervisados, que requieren datos etiquetados para su entrenamiento, o son técnicas que acompañan a los modelos, entre los más comunes se encuentran:

- Gradient boost.
- Bootstrapping.
- MaxEnt.
- Micro ARN (MIARN).
- Sistemas de Aprendizaje Significativo Integrado (SASI).
- Model-Key (local).
- Aprendizaje autónomo y “sumergirse” (SCUBA).
- Red Neuronal de Regresión General (GRNN).
- Redes neuronales convolucionales y máquinas de vectores de soporte (CNN-SVM).
- Algoritmo de Clasificación Basado en Lógica Difusa (FBCA).
- Librería de Máquinas de Vectores de Soporte (LibSVM).

- Redes neuronales convolucionales + memoria a largo y corto plazo + Redes neuronales profundas (CNN+LSTM+DNN).

Modelos avanzados:

- BERT (Bidirectional Encoder Representations from Transformers) es un modelo desarrollado por Google que comienza con un preentrenamiento no supervisado utilizando grandes volúmenes de texto sin etiquetas, como Wikipedia y libros. Este preentrenamiento le permite aprender representaciones del lenguaje. Luego, BERT se ajusta para tareas específicas como clasificación de texto, análisis de sentimientos y extracción de información, entre otras.
- Senti4SD, diseñada para analizar la polaridad emocional en canales de comunicación de desarrolladores, como preguntas y respuestas de Stack Overflow. Utiliza un enfoque de aprendizaje supervisado con documentos anotados como estándar para entrenar y validar el modelo. Clasifica el texto en categorías de polaridad como "positivo", "neutral" o "negativo", además, incluye una técnica para personalizar el clasificador con un estándar anotado por el usuario.
- SentiCR, herramienta o técnica de análisis de sentimientos diseñada para revisiones de código (code reviews). Su propósito principal es analizar comentarios de revisión de código para detectar el tono emocional o la polaridad (positivo, negativo, neutral) de las opiniones expresadas. Genera un vector de características basado en un modelo de "bolsa de palabras" y el algoritmo TF-IDF (Term Frequency-Inverse Document Frequency) es una técnica ampliamente utilizada en el campo del procesamiento del lenguaje natural (NLP) y la recuperación de información para medir la relevancia de una palabra dentro de un documento, considerando su aparición en un conjunto.

- CoreNLP es una herramienta de Stanford diseñada para generar anotaciones lingüísticas en textos, proporcionando funcionalidades como análisis gramatical, reconocimiento de entidades, análisis de dependencias y atribución de citas. Es muy utilizada en el desarrollo de modelos de aprendizaje automático y aplicaciones de análisis de texto. Mientras que herramientas como Senti4SD y SentiCR se especializan en análisis de sentimientos en contextos técnicos y de código, CoreNLP destaca por su versatilidad y aplicabilidad en diversos ámbitos del procesamiento de lenguaje natural, siendo un estándar en la investigación y desarrollo de sistemas NLP.

- EmoTxt tiene como objetivo identificar emociones específicas, como miedo o alegría, utilizando datos de plataformas como Stack Overflow y Jira, y está entrenada con un algoritmo de Máquinas de Soporte Vectorial (SVM). Su especialización en emociones concretas le permite destacarse frente a herramientas más generales de análisis de polaridad.

### ***Evaluación del Modelo***

La mayoría de los modelos utilizados en el análisis de sentimientos son de aprendizaje supervisado, y su objetivo principal es clasificar el texto en etiquetas predefinidas que reflejan el sentimiento expresado en los posts de una plataforma determinada. Para evaluar el rendimiento de estos modelos, se emplea un conjunto de datos de prueba y se calculan diversas métricas de evaluación que permiten medir su precisión y efectividad. Las definiciones de estas métricas se encuentran detalladas en los trabajos de Bishop, C. M. (2006) y Goodfellow, I., Bengio, Y., y Courville, A. (2016). Entre las métricas más comunes para clasificación se incluyen las siguientes:

- **Matriz de confusión:** Es una tabla que permite visualizar el rendimiento de un modelo de clasificación. Muestra los resultados de las predicciones comparadas con las clases

reales, permitiendo identificar los aciertos y los errores en las predicciones, verdaderos positivos (TP), falsos positivos (FP), verdaderos negativos (TN), falsos negativos (FN).

- **Exactitud (Accuracy):** Es la proporción de predicciones correctas sobre el total de predicciones. Se calcula como el número de predicciones correctas dividido entre el número total de predicciones. Es útil cuando las clases están equilibradas, pero puede ser engañosa en el caso de clases desbalanceadas.

$$Exactitud = \frac{TP + TN}{TP + FP + TN + FN}$$

- **Precisión (Precision):** Mide la proporción de verdaderos positivos entre los elementos que el modelo clasificó como positivos. Es útil cuando el costo de los falsos positivos es alto.

$$Precisión = \frac{TP}{TP + FP}$$

- **Recall (Sensibilidad o Tasa de Verdaderos Positivos):** Indica la proporción de verdaderos positivos identificados por el modelo con respecto al total de elementos que realmente son positivos. Es relevante cuando los falsos negativos son más costosos.

$$Recall = \frac{TP}{TP + FN}$$

- **F-Score:** Es la media armónica entre la precisión y el recall. Proporciona un balance entre ambas métricas, especialmente cuando hay una desproporción entre los falsos positivos y falsos negativos.

$$FScore = 2 * \frac{Precisión * Recall}{Precisión + Recall}$$

- **AUC-ROC (Área bajo la curva - Curva de Características Operativas del Receptor):** Mide la capacidad del modelo para distinguir entre clases positivas y negativas. AUC

(Área Bajo la Curva) es un valor entre 0 y 1, donde un valor cercano a 1 indica un modelo excelente.

- **Log Loss (Pérdida Logarítmica):** Mide la incertidumbre de las predicciones del modelo, penalizando las probabilidades incorrectas. Un valor bajo de Log Loss indica que el modelo está haciendo predicciones confiables.
- **Coefficiente de Gini:** Se utiliza para medir la desigualdad de una distribución. En el contexto de clasificación, se utiliza en modelos como los árboles de decisión y es una medida de pureza de los nodos. El coeficiente de Gini varía entre 0 (perfecto) y 1 (completamente impuro).

### ***Implementación y Monitoreo del Modelo***

Para maximizar el valor de un modelo, es fundamental implementarlo en un entorno de producción, donde pueda aplicarse a datos nuevos y generar un impacto real en la organización, proporcionando información útil y oportuna para la toma de decisiones. Las claves para una implementación exitosa son las siguientes:

- **Integración con sistemas existentes:** El modelo debe ser integrado de manera fluida en los sistemas operativos actuales, permitiendo que reciba entradas y entregue salidas de forma eficiente. Además, los usuarios deben poder interactuar con el modelo a través de interfaces intuitivas y accesibles.
- **Monitoreo continuo del rendimiento:** Una vez integrado, es crucial realizar un monitoreo constante del modelo para asegurarse de que sigue funcionando correctamente y entregando resultados precisos. Este seguimiento es esencial para detectar problemas de degradación en el rendimiento o cambios en los datos que puedan afectar la precisión del modelo.

- **Reentrenamiento y ajustes:** Con el tiempo, los modelos pueden requerir reentrenamiento o ajustes para adaptarse a nuevas tendencias o variaciones en los datos, lo que garantiza que continúe siendo relevante y preciso en su desempeño.

### Grupo 3

Análisis y resultados de la implementación de modelos de aprendizaje de máquina en la Red Social X, agrupando estudios y proyectos que aplican estos modelos. A través de una revisión exhaustiva, el objetivo es ofrecer una comprensión profunda de los efectos que estos modelos tienen en la interacción de los usuarios, así como en la dinámica social y política dentro de la plataforma. A continuación, se presenta una tabla con los ejemplos que se detallarán en esta sección.

**Tabla 3**

#### *Estudios e Investigaciones*

Autores	Nombre	Modelo usado	Resumen
Deng, H. et al. (2020)	Sentiment analysis of real-world migraine tweets for population research	Rsentiment	Analizar el texto de aquellas personas que comparten sus experiencias a través de posts en la red social X
Park, S. et al. (2023)	Mind games: A temporal sentiment analysis of the political messages of the Internet Research Agency on Facebook and Twitter	Rsentiment	Estudio de análisis de sentimiento sobre post y comentarios realizados en Facebook y X por IRA durante las elecciones presidenciales 2016 en EE. UU.
Sarsam S. et al. (2020)	Sarcasm detection using machine learning algorithms in Twitter: A systematic review	Modelos tradicionales y modelos adaptados	Comprender los significados literales y figurados es esencial para interpretar las opiniones de los usuarios en redes sociales, especialmente en la detección de sarcasmo. Este estudio revisa 31 investigaciones sobre algoritmos de aprendizaje automático para identificar sarcasmo en X, utilizando la metodología PRISMA

del Castillo Collazo, N. y Macías Herrera, R. M. (2024)	Trump vs Biden: Análisis de sentimientos de las publicaciones en español realizada en Twitter entre Estados de la Unión en las Elecciones Presidenciales del 2020	Modelos tradicionales y modelos adaptados	Análisis de sentimientos utilizando herramientas de procesamiento de lenguaje natural (PLN), que clasifican los textos en sentimientos positivos, negativos o neutros. Las publicaciones en Twitter fueron recolectadas durante un período relevante antes de las elecciones presidenciales de 2020
--	--	--	--

---

*Nota.* Estudios e Investigaciones sobre Modelos de Aprendizaje Automático Aplicados a X.

### ***Sentiment Analysis of Real-World Migraine Tweets for Population Research***

Deng, H. et al. (2020) analizan los posts sobre migrañas en la plataforma X, enfocándose en las experiencias de los usuarios y sus comportamientos sociales para futuras investigaciones sobre el impacto de la migraña. El método utilizado fue:

- **Criterios de Inclusión y Exclusión:** Se incluyeron posts que mencionaban "migraña" de usuarios con cuentas públicas que publicaron al menos un post entre el 6 de noviembre y el 7 de diciembre de 2014. Se excluyeron posts no relacionados, en otros idiomas distintos del inglés o de cuentas organizativas.
- **Recopilación de Datos:** Se obtuvieron datos sobre los posts y perfiles de usuario mediante la API de Twitter, procesándolos con R y Python.
- **Preprocesamiento:** Se eliminaron tweets con solo emojis o enlaces, así como datos geográficos incompletos.
- **Privacidad del Usuario:** Se protegió la información personal según las políticas de privacidad de X (Twitter en su momento).
- **Análisis Estadístico:** Se resumieron las características de los usuarios y se realizó un análisis multivariante para evaluar la actividad en diferentes grupos.

- **Procesamiento del Lenguaje Natural (PLN):** Se aplicaron técnicas como análisis de frecuencia de palabras y N-gramas para identificar patrones en los posts, aunque se reconocieron limitaciones en la comprensión contextual por parte de las máquinas.
- **Análisis del Sentimiento:** Los tweets fueron tokenizados y se calcularon puntuaciones de sentimiento utilizando el paquete sentiment en R.
- **Modelización de Temas:** Se utilizó Asignación de Dirichlet Latente (LDA) para identificar temas comunes en los posts, visualizados mediante nubes de palabras.
- Los resultados de este trabajo muestran que los usuarios mostraron patrones sentimentales variados, con publicaciones negativas más frecuentes que las positivas, reflejando frustración durante episodios de migraña. Los posts negativos a menudo incluían blasfemias, mientras que los positivos se centraban en el alivio. La clasificación errónea del sentimiento fue un desafío, especialmente con posts sarcásticos. Además, se identificaron dificultades en el modelado temático debido a la mezcla de temas en los datos ruidosos. Este estudio proporciona una base para investigar el impacto social y emocional de la migraña a través de las experiencias compartidas en redes sociales.

### ***Mind Games: A Temporal Sentiment Analysis of the Political Messages of the Internet***

#### ***Research Agency on Facebook and Twitter***

El artículo de Park, S. et al. (2023) investiga el uso de un modelo de análisis de sentimiento basado en lenguaje natural para analizar comentarios y publicaciones en Facebook y X de la Agencia de Investigación de Internet (IRA), conocida por su papel en la propaganda durante las elecciones presidenciales de EE. UU. en 2016. El estudio examina la dinámica emocional de los mensajes a través de tres dimensiones: plataformas sociales, identidad partidista y la identidad afroamericana.

Se utilizaron dos conjuntos de datos: 3,519 anuncios de Facebook asociados a la IRA y casi 3 millones de publicaciones en X. Para el análisis, se implementó el paquete Rsentiment, que desarrolla un modelo de procesamiento del lenguaje natural (PLN) para evaluar el sentimiento en los mensajes. Este modelo clasifica las emociones detectadas como positivas o negativas y asigna puntuaciones numéricas que indican el grado de sentimiento en cinco categorías, desde muy positivo hasta muy negativo. Rsentiment utiliza técnicas como el etiquetado de partes del discurso (Part of Speech tagging) para analizar cada palabra y también considera la aparición secuencial de negadores, lo que permite un análisis más preciso a nivel de frase.

Los resultados mostraron que Facebook tenía emociones predominantemente negativas antes de las elecciones, mientras que X mantuvo un sentimiento negativo constante. Además, los trolls que es el nombre dado aquellas personas reales que trabajan en IRA, quienes crean y difunden contenido diseñado para influir en la opinión pública y sembrar discordia, se dividen en los de izquierdas y derechas empleando estrategias emocionales opuestas en ambas plataformas, afectando potencialmente la participación política, especialmente entre votantes afroamericanos.

El estudio sugiere que la IRA utilizó una estrategia de "divide y vencerás", desestabilizando inicialmente el discurso público con mensajes negativos y luego reforzando identidades intergrupales a medida que se acercaban las elecciones, evidenciando una escalada emocional en los cinco meses previos a los comicios.

### ***Sarcasm Detection Using Machine Learning Algorithms in Twitter: a Systematic Review***

El reconocimiento de los significados literales y figurados es esencial para entender las opiniones de los usuarios sobre diversos temas en las redes sociales. En este contexto, el artículo de Sarsam, S. et al. (2020) presenta una revisión de 31 estudios utilizando la metodología

PRISMA, en la que se exploran diversos algoritmos de aprendizaje automático aplicados a la detección de sarcasmo en la plataforma de X. Los estudios fueron clasificados en dos grupos: Adapted Machine Learning Algorithms (AMLA) y Customized Machine Learning Algorithms (CMLA). El sarcasmo, como forma compleja de ironía, es frecuente en las interacciones en redes sociales, y detectar los posts sarcásticos es clave para la clasificación de textos. Este proceso tiene implicaciones significativas, ya que la correcta identificación del sarcasmo mejora la precisión en el análisis de sentimientos y opiniones. Los resultados de la revisión indicaron que el algoritmo SVM (máquinas de vectores de soporte) fue el más utilizado en el grupo AMLA para detectar sarcasmo en X. Además, se encontró que tanto SVM como CNN (redes neuronales convolucionales) ofrecieron un alto rendimiento en la predicción del sarcasmo.

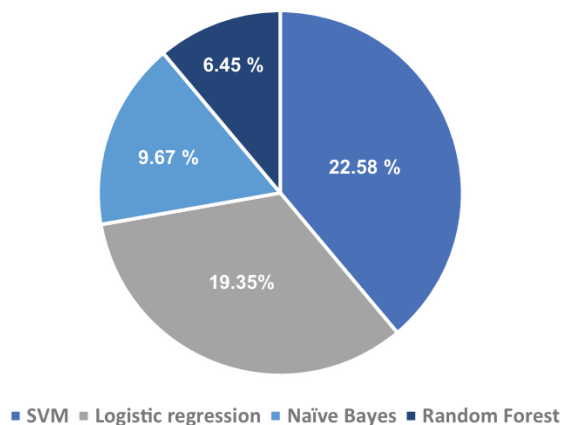
Por otro lado, los CMLA emplearon diferentes características de procesamiento de texto y clasificación, lo que generó variaciones en el desempeño de los algoritmos. De los resultados obtenidos, se concluyó que SVM y CNN-SVM fueron los algoritmos más eficientes para detectar sarcasmo en X. Se observó que las características léxicas, pragmáticas, de frecuencia y el etiquetado POS (partes del discurso) mejoraron el rendimiento de SVM. Además, las características léxicas y personales fueron los principales predictores del sarcasmo cuando se utilizó el algoritmo CNN-SVM. Con base en estos hallazgos, se recomienda incorporar características léxicas, pragmáticas, de frecuencia, etiquetado POS y atributos personales para mejorar el reconocimiento del sarcasmo en las plataformas de microbloggings como X.

Sin embargo, el estudio presenta ciertas limitaciones, como su enfoque exclusivo en la detección de sarcasmo en X, lo que deja de lado otras plataformas o tipos de ironía. Además, se excluyó el análisis de trabajos centrados en la detección de ironía general. Otros aspectos relacionados con los datos, como el número de palabras utilizadas, las técnicas de normalización

o parámetros clave de los algoritmos, como el tipo de núcleo, no fueron tratados en la revisión. En consecuencia, se sugiere que futuros estudios consideren estos factores, ya que podrían influir en el rendimiento de los algoritmos AMLA y CMLA en la detección de sarcasmo.

### Figura 3

*Frecuencia de Algoritmos AMLA*



*Nota.* Tomado de Sarsam, S., et al. (2020) “The frequency of AMLA’s group algorithms”.

### ***Trump vs Biden: Análisis de Sentimientos Elecciones Presidenciales del 2020***

En base a un estudio realizado por Del Castillo Collazo, N. y Macías Herrera, R. M. (2024) en la Universidad Santo Tomas, se realizó el análisis de sentimientos utilizando herramientas de procesamiento de lenguaje natural (PLN), que clasifican los textos en sentimientos positivos, negativos o neutros. Las publicaciones en X fueron recolectadas durante un período relevante antes de las elecciones presidenciales de 2020. El modelo utilizado fue el léxico NRC (National Research Council), que se asocia con ocho emociones: alegría, asombro, confianza, disgusto, ira, miedo, premonición y tristeza, así como los sentimientos de positivo y negativo (Saif & Peter, 2013).

- Fuente de datos: Extracción de los posts que mencionan a los candidatos Trump y Biden, filtrados por hashtags y palabras clave relacionadas con las elecciones presidenciales de 2020, como #Trump2020, #BidenHarris, #Elecciones2020, entre otras. Se empleó una API de X (Twitter para su época) para recopilar información pública geolocalizada y en idioma español, específicamente en los estados seleccionados de Florida, California y Texas (del Castillo Collazo, N. y Macías Herrera, R. M. 2024).

#### Distribución del sentimiento

- Trump: Un porcentaje significativo de tweets mostró sentimientos negativos (55%), relacionados con temas como políticas migratorias y su manejo de la pandemia. Sin embargo, también hubo un segmento considerable de apoyo positivo (30%), reflejando el respaldo de ideologías conservadoras.

- Biden: Los tweets a favor de Biden mostraron una mayor proporción de sentimientos positivos, destacando temas como la inclusión y el enfoque en la recuperación económica. Los sentimientos negativos hacia Biden estuvieron relacionados con dudas sobre su liderazgo y ataques de opositores.

- Sentimientos neutros: Reflejaron publicaciones informativas o análisis objetivos, sin inclinaciones emocionales claras. Este segmento es clave para identificar el impacto de la desinformación y la neutralidad en redes sociales.

#### Relevancia de los Estados Analizados

- Florida: Importante por su alto número de votantes hispanos de ascendencia cubana y venezolana.

- Texas: Un bastión tradicionalmente republicano con un creciente electorado latino diverso.

- California: Estado con una alta densidad de votantes latinos y una influencia creciente en las elecciones. Estos tres estados fueron seleccionados debido a su relevancia electoral y densidad poblacional hispana.

Las redes sociales desempeñan un papel fundamental en la movilización del voto, influyendo en percepciones y generando entusiasmo entre los votantes indecisos. Los posts en español permiten una conexión directa con comunidades que pueden afectar los resultados electorales. Durante las elecciones presidenciales de 2020, Twitter se destacó como una herramienta clave para formar opiniones entre jóvenes votantes y facilitar diálogos políticos en tiempo real. Los candidatos aprovecharon estas plataformas para comunicarse directamente con electores hispanohablantes, destacando la importancia de las redes sociales en el proceso electoral.

**Información Positiva:** Los usuarios compartieron propuestas de ambos candidatos, análisis de debates y mensajes motivacionales dirigidos a la comunidad latina.

**Desinformación:** Se identifican campañas de noticias falsas que buscaban polarizar al electorado hispano, muchas veces utilizando temas sensibles como la inmigración y la economía. Esto destaca la necesidad de filtrar y validar el contenido en redes (Arias, M. 2017).

## Conclusiones

Es fundamental reconocer que la relación e influencia en plataformas como X trasciende la mera interacción entre figuras públicas y su audiencia; la propia plataforma desempeña un papel crucial en cómo se establece esta influencia, afectando a ambos grupos. Cada red social presenta un desarrollo único que varía según las características culturales y sociales de los países en los que opera, lo que merece un análisis más profundo sobre cómo estas particularidades influyen en la generación y distribución de información.

Los estudios revisados y el análisis del impacto de la red social X revelan una creciente influencia de las plataformas sociales en la gobernanza estatal. Estas herramientas se han convertido en espacios clave para la interacción entre figuras públicas, líderes políticos y ciudadanos, modelando la percepción pública de las decisiones gubernamentales. La comunicación instantánea y masiva en redes como X no solo facilita la participación ciudadana y la toma de decisiones en tiempo real, sino que también genera patrones emergentes de polarización, impactando la legitimidad de las políticas estatales y la dinámica política en general. Por lo tanto, es crucial comprender cómo estas plataformas evolucionan y amplían su rol en la gobernanza, ya que pueden moldear la agenda política y la relación entre gobierno y sociedad.

Las herramientas utilizadas para el análisis de sentimiento y clasificación de texto no garantizan la cobertura completa de todos los posts en X ni aseguran que los datos recopilados sean aleatorios. Es importante considerar que la información obtenida puede estar sujeta a sesgos, lo que podría afectar los resultados del análisis.

Distinguir entre interacciones humanas y actividades de bots en plataformas como X representa un desafío significativo. Esta dificultad puede comprometer la confiabilidad de los resultados, ya que los bots pueden distorsionar el análisis de sentimientos y opiniones.

El uso de modelos de análisis de sentimiento no se limita a identificar y clasificar texto; estas clasificaciones son insumos para modelos más complejos. Por ello, es crucial definir claramente el objetivo del modelo y el propósito de los resultados dentro del contexto más amplio del análisis en X.

La revisión de estudios sobre el uso de modelos de aprendizaje automático para el análisis de sentimientos y clasificación de textos en X destaca la efectividad de metodologías como redes neuronales profundas y algoritmos de procesamiento de lenguaje natural (PLN) para comprender las opiniones políticas y sociales en tiempo real. Los resultados de diversos proyectos muestran cómo la detección de emociones y tendencias políticas, ya sean positivas, negativas o neutras, puede predecir cambios en la percepción social y su influencia en la política. Este análisis permite estudiar la diseminación de emociones y opiniones en las redes sociales, contribuyendo a una mejor comprensión de los comportamientos electorales y el análisis político-social.

## Recomendaciones

Según los resultados de la investigación presentada en esta monografía, se recomienda el uso de técnicas de aprendizaje automático, especialmente los modelos de Support Vector Machine (SVM), Redes Neuronales Convolucionales combinadas con SVM (CNN-SVM) y Bidirectional Encoder Representations from Transformers (BERT), debido a su destacado rendimiento, precisión y capacidad para obtener resultados óptimos en el análisis de sentimientos y clasificación de texto. No obstante, es importante señalar que la eficiencia de estos modelos puede variar dependiendo de las características específicas de cada proyecto.

Estudiar y analizar la evolución de la participación política digital es fundamental para entender cómo han cambiado las plataformas digitales y las estrategias políticas en los últimos años. Este estudio debe abarcar tanto los avances tecnológicos como sus efectos en la participación política y la relación entre políticos y ciudadanos. La rápida adaptación de las plataformas a nuevas necesidades sociales y políticas requiere una reflexión constante sobre sus implicaciones en la esfera pública.

Evaluar la capacidad crítica de los usuarios es crucial, e investigar cómo los ciudadanos filtran y evalúan la información política en redes sociales. Se debe analizar el impacto de la educación digital, la alfabetización mediática y las competencias críticas en la interpretación de mensajes políticos. También es importante estudiar la conciencia de los usuarios sobre estrategias de manipulación de datos y segmentación, y cómo estos factores afectan la calidad de sus opiniones.

Las estrategias políticas han evolucionado desde la difusión unidireccional hasta interacciones más complejas, como campañas personalizadas. Se recomienda investigar cómo los políticos adaptan sus mensajes a cada plataforma, aprovechando la segmentación de audiencias y

el análisis de datos. Además, es esencial identificar las diferencias y similitudes entre la interacción política en redes sociales y métodos tradicionales para comprender mejor su impacto en la política.

Es importante estudiar cómo las plataformas influyen en la relación entre políticos y ciudadanos, más allá de su papel como canales de comunicación. Las redes sociales actúan como actores activos que filtran, amplifican y distorsionan mensajes. Se debe analizar cómo los algoritmos, la personalización del contenido y las políticas de moderación afectan la percepción pública de figuras políticas y sus propuestas.

Dada la rápida difusión de contenido viral, es imprescindible investigar cómo la desinformación afecta la participación política. Esto incluye el análisis del consumo y respuesta a noticias falsas, así como el impacto de discursos de odio en el debate público y en procesos democráticos, especialmente durante elecciones.

La recolección y uso de datos personales para fines políticos plantea dilemas éticos significativos. Es necesario investigar las implicaciones del uso masivo de datos para crear perfiles de votantes y dirigir campañas personalizadas. Los estudios deben enfocarse en la privacidad, el consentimiento informado y la transparencia para mantener la confianza pública en el sistema político digital, así como explorar políticas regulatorias que equilibren innovación tecnológica con protección de derechos.

Dado el impacto significativo de la red social X en la gobernanza estatal y su capacidad para moldear la opinión pública, se recomienda a los gobiernos y actores políticos integrar de manera más efectiva el análisis de sentimientos y la clasificación de textos en sus estrategias de comunicación y toma de decisiones.

## Referencias

- Akkaya, C., Wiebe, J. y Mihalcea, R. (2009). *Desambiguación del sentido de las palabras a partir de la subjetividad. Actas de la Conferencia de 2009 sobre métodos empíricos en el procesamiento del lenguaje natural*, 190–199. <https://doi.org/10.3115/1699510.1699540>
- Arias, M. D. (2017). *La posverdad y las fake news: Un análisis desde la comunicación política*. <https://orcid.org/0000-0002-1736-4797>
- Arroyabe (s. f.). *Media Landscapes Expert Analyses of the media, COLOMBIA*. <https://medialandscapes.org/country/colombia/media/social-networks>.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Editorial Springer.
- Bollen, J., Mao, H. y Zeng, X. (2011). *El estado de ánimo en Twitter predice el mercado de valores*. *Journal of Computational Science*, 2(1), 1–8. <https://doi.org/10.1016/j.jocs.2010.12.007>
- Bossetta, M. (2018). *The Digital Architectures of social media: Comparing Political Campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 U.S. Election*. *Journalism & Mass Communication Quarterly*, 95(2), 471-496. <https://doi.org/10.1177/1077699018763307>
- Boucher, JD. y Osgood, CE (1969). *La hipótesis de Pollyanna*. *Journal of Abnormal and Social Psychology*, 62(3), 350-359.
- Boulianne, S. y Larsson, A. O. (2023). *Engagement with candidate posts on Twitter, Instagram, and Facebook during the 2019 election*. *New Media & Society*, 25(1), 119-140. <https://doi.org/10.1177/14614448211009504>

- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Shinn, J., Wu, J. y Amodei, D. (2020). *Language models are few-shot learners. Proceedings of NeurIPS 2020*. <https://arxiv.org/abs/2005.14165>
- Cambria E. (2017). "Affective Computing and Sentiment Analysis," in *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 102-107, Mar.-Apr. 2016, doi: 10.1109/MIS.2016.31. <https://ieeexplore.ieee.org/document/7435182>
- Colprensa (2023). *¿Hiperconectados? Cuatro de cada diez colombianos gastan su tiempo libre en redes sociales*. Periódico: Diario El País. <https://www.elpais.com.co/colombia/hiperconectados-cuatro-de-cada-diez-nos-gasta-su-tiempo-libre-en-redes-sociales.html>
- Criado, J. y Villodre, J. (2018). *Comunicando datos masivos del sector público local en redes sociales. Análisis de sentimiento en Twitter. Profesional de la información*. Information Professional, 27(3), 624-632.
- Damasio, A. (2007). *En busca de Spinoza: Neurobiología de la emoción y los sentimientos*. Editorial Planeta.
- Del Castillo Collazo, N. y Macías Herrera, R. M. (2024). *Trump vs Biden: Análisis de sentimientos de las publicaciones en español realizada en Twitter en tres Estados de la Unión en las Elecciones Presidenciales del 2020*. Comunicaciones En Estadística, 1(17). <https://doi.org/10.15332/23393076.10154>
- Deng, H., Wang, Q., Turner, D., Sexton, K., Burns, S., Eikermann, M., Liu, D., Cheng, D., y Houle, T. (2020). *Sentiment analysis of real-world migraine tweets for population research. Cephalalgia Reports*. <https://doi.org/10.1177/2515816319898867>
- Deuze, M. (2012). *Vida mediática*. Prensa política.

- Devlin, J., Chang, M., Lee, K., Toutanova, K. (2018) "*BERT: Pre-training of deep bidirectional transformers for language understanding.*" arXiv preprint arXiv:1810.04805.  
<https://arxiv.org/abs/1810.04805>
- Documentation home (2024). *Docs | Twitter Developer Platform.*  
<https://developer.x.com/en/docs>
- El, F., Ferrara, E., y Flammini, A. (2015). *Los bots sociales distorsionan el debate en línea sobre las elecciones presidenciales estadounidenses de 2016.* First Monday, 20 (9) .  
<https://doi.org/10.5210/fm.v20i9.6417>
- Es, M. (2024). *Twitter ahora admite oficialmente el contenido sexual: cambios en X, límites y preocupación.* Maldita.es - Periodismo para que no te la cuelen.  
<https://maldita.es/malditatecnologia/20240607/x-twitter-permite-contenido-sexual/>
- Fernández, E. (2024). *Conoce la historia del machine learning: desde sus inicios Tokio School.*  
<https://www.tokioschool.com/noticias/historia-machine-learning/>
- García, D., Garas, A., y Schweitzer, F. (2012). *El sesgo positivo de la difusión de la información.* PLoS ONE, 7(10), e47276.
- Goodfellow, I., Bengio, Y. y Courville, A. (2016). *Deep Learning.* MIT Press.
- Hootsuite Inc. (2024). *Social Media Marketing & Management Dashboard - Hootsuite.* Social Media Marketing and Management Tool | Hootsuite. <https://www.hootsuite.com/>
- Honnibal, M. y Montani, I. (2017). *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks, and incremental parsing.* To appear in Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018). <https://spacy.io/>

- Hutchinson, A. (2022). *New Study Shows Twitter is the Most Used Social Media Platform Among Journalists*. *Social Media Today*. <https://www.socialmediatoday.com/news/new-study-shows-twitter-is-the-most-used-social-media-platform-among-journa/626245/>
- Hutto, CJ y Gilbert, E. (2014). *VADER: Un modelo parsimonioso basado en reglas para el análisis de sentimientos de textos en redes sociales*. Actas de la 8.<sup>a</sup> Conferencia Internacional sobre Weblogs y Redes Sociales (ICWSM) , 216–225. <https://doi.org/10>
- Jurafsky & Martin (2019). *Speech and Language Processing*.  
<https://web.stanford.edu/~jurafsky/slp3/>
- Kuss, D. J., & Griffiths, M. D. (2017). *Social networking sites and addiction: Ten lessons learned*. *International Journal of Environmental Research and Public Health*, 14(3), 311. <https://doi.org/10.3390/ijerph14030311>
- Kwak, H., Lee, C., Park, H., y Moon, S. (2010). *¿Qué es Twitter, una red social o un medio de noticias?*. Actas de la 19<sup>a</sup> Conferencia Internacional sobre la World Wide Web, 591-600.
- Lasri, I. & Riadsolh, A. y Elbelkacemi, M. (2023). *Real-time Twitter Sentiment Analysis for Moroccan Universities using Machine Learning and Big Data Technologies*. *International Journal of Emerging Technologies in Learning (Ijet)*, 18(05), 42-61. <https://doi.org/10.3991/ijet.v18i05.35959>
- Loria, S. (2018). *TextBlob: Simplified text processing*. <https://textblob.readthedocs.io/en/dev/>
- Mikolov, T., Chen, K., Corrado, G., y Dean, J. (2013). *Efficient estimation of word representations in vector space*. *Proceedings of the International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1301.3781>

- Moreno V., Camacho, F. y Ramirez, A. (2018). *Análisis de Sentimientos en Redes Sociales (Twitter) Mediante Python para la Detección de Oportunidades de Negocio*. TecnoCultura, 45. <https://tecnocultura.org/index.php/Tecnocultura/article/view/175>
- Park, S., Strover, S., Choi, J. y Schnell, M. (2023). *Mind games: A temporal sentiment analysis of the political messages of the Internet Research Agency on Facebook and Twitter*. *New Media & Society*, 25(3), 463-484. <https://doi.org/10.1177/14614448211014355>
- Pennington, J., Socher, R. y Manning, C. D. (2014). *GloVe: Global vectors for word representation*. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1532–1543. <https://aclanthology.org/D14-1162/>
- PRISMA Statement. (2024). *History and development*. <https://www.prisma-statement.org/history-and-development>
- Radford, A., Narasimhan, K., Salimans, T., y Sutskever, I. (2018). *Improving language understanding by generative pre-training*. OpenAI Blog. <https://openai.com/research/language-unsupervised>
- Robertson, C. (2023). *Here's what our research says about news audiences on Twitter, the platform now known as X*. Reuters Institute for The Study of Journalism. University of Oxford. <https://reutersinstitute.politics.ox.ac.uk/news/heres-what-our-research-says-about-news-audiences-twitter-platform-now-known-x>
- Sahly, A. & Shao, C. y Kwon, K. (2019). *Social Media for Political Campaigns: An Examination of Trump's and Clinton's Frame Building and Its Effect on Audience Engagement*. *Social media + Society*, 5(2). <https://doi.org/10.1177/2056305119855141>.
- Saif, H. y Peter, D. (2013). *Léxico de emociones de la NRC . Consejo Nacional de Investigaciones*. <http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>

Sanhueza, A. (2013). *Formas de representación de textos cortos para su posterior clasificación automática* [Tesis de grado, Pontificia Universidad Católica de Valparaíso].

[http://opac.pucv.cl/pucv\\_txt/txt-3000/UCE3026\\_01.pdf](http://opac.pucv.cl/pucv_txt/txt-3000/UCE3026_01.pdf).

Silberling, A., Corral, C., y Stringer, A. (2024). *Elon Musk's X: A complete timeline of what Twitter has become*. TechCrunch. <https://techcrunch.com/2024/06/05/elon-musk-twitter-everything-you-need-to-know/>

Sprout Social. (2024). *Sprout social pricing - How much does Sprout Social cost?*

<https://sproutsocial.com/pricing/>

Sunstein, C. (2018). *Cámaras de eco: el estudio científico de la polarización*. Prensa de la Universidad de Cambridge.

Trump vs Biden (2020). *Análisis de sentimientos de publicaciones en Twitter durante las elecciones presidenciales de 2020*. <https://doi.org/10.15332/23393076.10154>

Van Vliet, L., Törnberg, P. y Uitermark, J. (2020). *The Twitter parliamentarian database: Analyzing Twitter politics across 26 countries*. PloS one, 15(9), e0237073.

<https://doi.org/10.1371/journal.pone.0237073>

We Are Social. (2024). *Digital 2024: 5 billion social media users. We Are Social Australia*.

<https://wearesocial.com/au/blog/2024/01/digital-2024-5-billion-social-media-users/>

Wike R., Silver L., Fetterolf J., Huang C., Austin S., Clancy L. y Gubbala S. (2022). *Social media Seen as Mostly Good for Democracy Across Many Nations, But U.S. is a Major Outlier* | Pew Research Center. Pew Research Center's Global Attitudes Project.

<https://www.pewresearch.org/global/2022/12/06/social-media-seen-as-mostly-good-for-democracy-across-many-nations-but-u-s-is-a-major-outlier/>

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A. y others. (2020). *Hugging Face Transformers, Transformers: State-of-the-art natural language processing.*

<https://huggingface.co/transformers/>

Zajonc, RB (1980). *Sentir y pensar: las preferencias no necesitan inferencias.* Psicólogo americano.

Zeynep Meral K., Dr. Yurdagül Meral Ü. (2021). *The Role of social media In Arab Spring.*

Research Gate DOI:10.17932/IAU.EJNM.25480200.2021/ejnm\_v5i1003