

Sentimientos hacia la vacunación contra la covid-19: panorama colombiano en Twitter

Augusto Rodríguez-Orejuela¹
Claudia Lorena Montes-Mora²
Carlos Fernando Osorio-Andrade³

Recibido: 01/07/2021
Aprobado por pares: 05/11/2021

Enviado a pares: 12/07/2021
Aceptado: 17/12/2021

DOI: 10.5294/pacla.2022.25.1.4

Para citar este artículo / to reference this article / para citar este artigo

Rodríguez-Orejuela, A., Montes-Mora, C. L. y Osorio-Andrade, C. F. (2022). Sentimientos hacia el vacunación contra la covid-19: panorama colombiano en Twitter. *Palabra Clave*, 25(1), e2514. <https://doi.org/10.5294/pacla.2022.25.1.4>

Resumen

El propósito de este documento consiste en analizar los sentimientos subyacentes en publicaciones de Twitter sobre la vacunación contra la covid-19. Para cumplir el objetivo, se extraen, mediante minería de datos, 38.034 publicaciones de esta red social y se aplican técnicas de *Machine Learning*, en concreto, análisis de sentimientos y análisis de redes, para identificar los sentimientos que expresan los usuarios de esta red social hacia la vacunación por covid-19. También se identifican las cuentas más importantes de Twitter en temas de vacunación. Los resultados sugieren que, en su mayoría, los sentimientos hacia las vacunas son negativos. El miedo y la ira, respectivamente, son las emociones más recurrentes en las opiniones de Twitter. Por otra parte, se identifica que las cuentas más relevantes pertenecen a medios de comunicación, políticos e influenciadores, los cuales se clasifican

1 <https://orcid.org/0000-0003-2865-1748>. Universidad del Valle, Colombia. augusto.rodriguez@correounivalle.edu.co

2 <https://orcid.org/0000-0002-8154-9737>. Universidad del Valle, Colombia. claudia.montes@correounivalle.edu.co

3 ✉ <https://orcid.org/0000-0002-5095-4991>. Universidad del Valle, Colombia.
carlos.fernando.osorio@correounivalle.edu.co

de acuerdo con los principales sentimientos respecto de la vacuna. Destaca la oposición al gobierno, con sentimientos de ira, y a medios de comunicación reconocidos, con emociones asociadas a la alegría.

Palabras clave (Fuente: tesaurus de la Unesco)

Análisis de sentimientos; covid-19; red social; Twitter; vacunas.

Feelings towards COVID-19 Vaccination: Colombian Panorama on Twitter

Abstract

This document intends to analyze the sentiments underlying COVID-19 vaccination tweets. To achieve the objective, 38,034 publications from this social network are extracted through data mining, applying Machine Learning techniques, specifically sentiment analysis and network analysis, to identify the feelings expressed by Twitter users. We also identify the most relevant Twitter accounts on vaccination issues. The results suggest that feelings about vaccines are primarily negative; fear and anger, respectively, are the most recurring emotions in Twitter opinions. Moreover, we noted that the most relevant accounts belong to the media, politicians, and influencers, classified according to their feelings toward the vaccine. Opposition to the government with feelings of anger and opposition to recognized media with joyful emotions stand out.

Keywords (Source: Unesco Thesaurus)

Sentiment analysis; COVID-19; social network; Twitter; vaccines.

Sentimentos sobre a vacina contra a covid-19: panorama colombiano no Twitter

Resumo

O objetivo deste artigo é analisar os sentimentos subjacentes em publicações do Twitter sobre a vacinação contra a covid-19. Para atingi-lo, são extraídas, mediante mineração de dados, 38.034 publicações dessa rede social e são aplicadas técnicas Machine Learning, em concreto, análise de sentimentos e análise de redes, para identificar os sentimentos que os usuários dessa rede expressam quanto à vacinação contra a covid-19. Também são identificadas as contas mais importantes do Twitter em temas de vacinação. Os resultados sugerem que, em sua maioria, os sentimentos sobre as vacinas sejam negativos. O medo e a ira, respectivamente, são as emoções mais recorrentes nas opiniões do Twitter. Por sua vez, é identificado que as contas mais relevantes pertencem a meios de comunicação, políticos e influenciadores, os quais são classificados de acordo com os principais sentimentos a respeito da vacina. Destaca-se a oposição ao governo, com sentimentos de ira, e a meios de comunicação reconhecidos, com emoções associadas à alegria.

Palavras-chave (Fonte: tesouro da Unesco)

Análise de sentimentos; covid-19; rede social; Twitter; vacinas.

Desde finales de 2019, la pandemia de coronavirus o covid-19 que surgió en Wuhan, China, se ha convertido en un grave problema de salud pública debido a la velocidad de transmisión que presenta (Crokidakis, 2020; Mukherjee, 2020). Los gobiernos alrededor del mundo han implementado múltiples medidas para controlar la propagación del virus y mitigar los efectos negativos de la covid sobre la población, adoptando el confinamiento obligatorio, el cierre de escuelas y de fronteras, el uso generalizado del cubrebocas y distanciamiento social (Atalan, 2020; Musinguzi y Asamoah, 2020). No obstante, y hasta el momento, aunque estas medidas han ayudado a evitar el colapso de los sistemas de salud, no han sido suficientes para retornar a la tan ansiada “normalidad” (Kuckertz *et al.*, 2020).

Actualmente, una de las esperanzas más grandes que tiene la humanidad para detener la pandemia de covid-19 reposa en la inmunización de la población a través de las vacunas que se han desarrollado a una velocidad sin precedentes (Stephoe y Francourt, 2021). Diversos investigadores científicos, gubernamentales y privados han trabajado de manera conjunta para desarrollar una vacuna probada que pueda proteger a las poblaciones del virus, al mismo tiempo que se salvaguarde la economía, debido a la interrupción que generan las medidas de confinamiento (Murphy *et al.*, 2021).

El 3 de marzo de 2020, en los Estados Unidos, tuvo lugar el primer ensayo clínico en humanos de una vacuna contra la covid-19 (Worldometer, 2020) y, poco tiempo después, se presentaron varios ensayos alrededor del mundo (OMS, 2020). Para el 11 de septiembre de 2020, un total de ocho vacunas habían avanzado a ensayos clínicos de fase 3 y 2, y recibido la aprobación para su uso temprano o limitado (The New York Times, 2020a). Para enero de 2021, algunas vacunas covid-19 habían demostrado su valor y obtenido autorizaciones de uso de emergencia en varios países; sin embargo, un grave problema parecía surgir en el horizonte y era la desconfianza de algunos sectores de la población en estos mecanismos de prevención (Murphy *et al.*, 2021). En un estudio conducido por Neumann *et al.* (2020), 7.664 personas residentes en siete países europeos (Francia, Alemania, Portugal, Italia, Países Bajos, Reino Unido y Dinamarca) fueron entrevistadas en relación con la vacuna contra la covid-19. El 18,9%

de los encuestados señalaron que no estaban “seguros” de recibir la vacuna, mientras que un 7,2% fue más enfático e indicó que no deseaba vacunarse. En España, el Centro de Investigaciones Sociológicas (CIS) indicó que en el 2020 solo una tercera parte de los españoles estaban dispuestos a recibir la vacuna; en Rusia, tan solo un 9% de la población parecía estar de acuerdo con esta medida de protección ante el virus (Suárez, 2020) y en México el 25% de los ciudadanos aseguró que no se vacunaría contra el coronavirus (Ortiz, 2020).

Hoy en día, las redes sociales son un escenario importante a través del cual se difunde información y desinformación sobre las vacunas; por lo tanto, estudiar las conversaciones y comentarios relacionados con las vacunas en estos medios podría contribuir a identificar elementos que contribuyan a la confianza en la vacunación y a desarrollar estrategias de comunicación y concientización, para disminuir el efecto de las noticias falsas sobre las vacunas (Piedrahíta *et al.*, 2021). Con el objetivo de analizar los sentimientos subyacentes en los comentarios de las redes sociales relacionados con la vacunación por covid-19 y, al mismo tiempo, identificar los actores más importantes frente a esta temática, en el presente estudio se desarrolla un análisis de sentimientos mediante *Machine Learning* a 38.043 trinos o tuits de usuarios colombianos difundidos en la red social Twitter, desde el 12 de mayo hasta el 30 de septiembre de 2020. Además, se aplican técnicas de grafos para identificar los usuarios más relevantes que difunden opiniones sobre la vacunación en Colombia y el sentimiento predominante que presentan.

Vale la pena señalar que en bases de datos indexadas, como Scopus, se registran algunas investigaciones que aplican el análisis de sentimientos y las técnicas de procesamiento de textos para analizar la opinión pública digital respecto a la vacunación, Piedrahíta *et al.* (2021) desarrollaron un sistema de clasificación de sentimientos en Twitter hacia las vacunas e identificaron que en 2019 el 69,3% de los comentarios fueron neutrales, el 21,78% positivos y el 8,86% negativos. Por otro lado, Sued (2020), usando un algoritmo para analizar la desinformación relacionada con la vacunación en videos de YouTube, identificó que algunos influenciadores promueven este tipo de opiniones, por lo que se ha difundido una red de videos de-

dicados a despotricar sobre las vacunas covid-19. Dubey (2021) también adelantó un análisis de sentimientos de trinos publicados en India con respecto a las vacunas covid-19 y encontró que, si bien la mayoría de las publicaciones demuestran sentimientos positivos hacia la vacunación, también se presentan sentimientos negativos asociados con esta medida, principalmente relacionados con emociones de miedo e ira.

A pesar de estas y otras investigaciones que aplican el análisis de sentimientos y las técnicas de clasificación de textos para analizar opiniones de redes sociales respecto a la vacunación por covid-19, hasta el momento, las investigaciones en América Latina son escasas, sin que existan en bases de datos indexadas estudios de este tipo para Colombia. Por lo tanto, el presente artículo atiende a esta brecha de investigación y espera servir como insumo para conocer las opiniones y sentimientos de los colombianos hacia las vacunas contra la covid-19. De este modo, la presente investigación se divide en seis apartados, incluida la introducción, seguida del marco referencial, la metodología empleada, los resultados del estudio, la discusión y, finalmente, las conclusiones más importantes.

Marco referencial

Resistencia a la vacunación

Los movimientos antivacunas son tan antiguos como las propias vacunas (Wolfe, 2002). En un principio, las primeras vacunas presentaban efectos adversos relevantes, incluso, en ocasiones, se evidenciaron decesos como consecuencia de la vacunación, lo que poco a poco dio lugar al surgimiento de grupos de detractores de las vacunas que se basaban en los riesgos de la vacunación, sin tener en cuenta los beneficios relacionados con la prevención de enfermedades (Salleras, 2018). En la actualidad, son muchos los argumentos que esgrimen los antivacunas para sustentar su posición ante la inoculación; entre los más importantes destacan: 1) argumentos filosóficos-religiosos, que suponen que las vacunas rompen el equilibrio de la naturaleza; 2) argumentos relacionados con la falta de eficacia, donde se sostiene que las enfermedades han disminuido, por los avances socioeconómicos y la inmunidad colectiva, y no por la vacunación; 3) razones rela-

cionadas con los riesgos y consecuencias de la vacunación; 4) argumentos propios de la corriente ortodoxa de la medicina homeopática, que asegura que las vacunas alteran la predisposición genética del individuo a desarrollar y padecer ciertas enfermedades; y 5) teorías conspirativas que presentan la vacunación como un plan sistemático de los gobiernos para hacer daño a las personas (Fernández y Baquero, 2019).

Más allá de las razones que llevan a ciertos grupos de la población a oponerse a la vacunación, lo que más preocupa a los organismos de salud alrededor del mundo son los problemas de salud pública relacionados con estos movimientos (Aps *et al.*, 2018). Quizás una de las consecuencias más notorias de la resistencia a la vacunación es la reaparición de brotes de enfermedades prevenibles, como ha ocurrido con el sarampión, la tos ferina y la influenza en el continente europeo y Estados Unidos (Carrasco y Lozano, 2018; Hotez, 2019; Zucker *et al.*, 2020). Sin embargo, con la llegada de la covid-19 y ante la urgencia de vacunar a la población global para frenar el contagio del virus, los movimientos antivacunas se hacen cada vez más peligrosos (Trogen *et al.*, 2020).

Redes sociales, noticias falsas y vacunación

Como se ha mencionado previamente, en la actualidad quizás la mayor esperanza que tiene el mundo para detener la enfermedad de covid-19 se encuentra en la difusión rápida de las vacunas que han demostrado su efectividad (Stephoe y Francourt, 2021). No obstante, en un momento donde se debería aplaudir el avance de la ciencia, las voces de los antivacunas han encontrado un nuevo aliento, reforzadas por la disminución de la confianza en los gobernantes, el impacto de las noticias falsas y, por supuesto, el efecto amplificador de las redes sociales (Ashton, 2021). Justamente, diversas investigaciones han señalado que las redes sociales contribuyen al aumento en la resistencia hacia las vacunas, ya que posibilitan la rápida difusión de rumores y mitos sobre la vacunación (European Centre for Disease Prevention and Control, 2020; Piedrahíta *et al.*, 2021). Hussein *et al.* (2020), por ejemplo, identificaron que las noticias falsas (*fake news*) relacionadas con la vacunación se encuentran dentro de las cinco teorías conspirativas más populares, junto con la falsa teoría de que el hombre no llegó

a la luna, la creencia terraplanista, la conspiración en torno al ataque a las Torres Gemelas y los senderos químicos, donde se sugiere que los aviones al volar liberan deliberadamente agentes químicos para afectar la población (Sued, 2020). Asimismo, Carrieri *et al.* (2019) analizan las noticias falsas en internet relacionadas con la vacunación (como la incidencia de estas en el autismo en niños) y concluyen que estas noticias son un importante impulsor de la no inmunización de la población. Finalmente, en plataformas de opinión como Twitter se han desarrollado múltiples estudios que analizan la difusión de información errónea sobre la vacunación, mediante técnicas de análisis de contenido, análisis de sentimientos y minería de textos (Jamison *et al.*, 2020; Piedrahíta *et al.*, 2021; Dubey, 2021). La importancia que tienen entonces las redes sociales en la percepción de la población respecto de las vacunas no debe pasar desapercibida. Las investigaciones que ayudan a entender los sentimientos de los usuarios de estos sitios frente a la vacunación pueden contribuir a la generación de estrategias de concientización que busquen disminuir los efectos de las noticias falsas y los rumores sobre la inmunización.

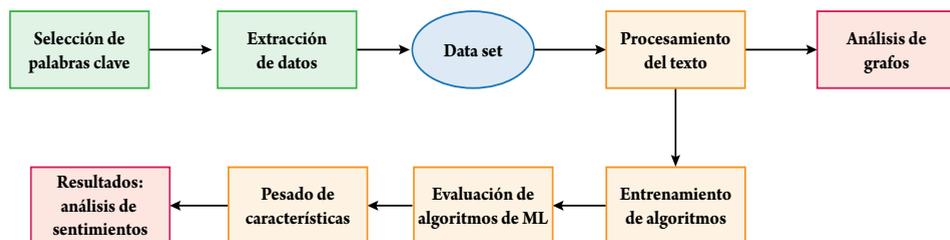
Metodología de investigación

Análisis de sentimientos en redes sociales

El análisis de sentimientos, también conocido como minería de opinión o análisis de emociones, consiste en la aplicación de técnicas computacionales o de *Machine Learning* para analizar opiniones, sentimientos y emociones expresadas en un cuerpo de texto, como noticias internacionales publicadas en la red o comentarios de personas en redes sociales (Liu, 2010). Desde el año 2001, el análisis de sentimientos se ha popularizado de manera considerable en la investigación científica y gerencial debido a los beneficios sustanciales que ofrece, ya que permite el análisis automático de grandes cantidades de texto que no se podrían estudiar de manera manual (Manning *et al.*, 2008). En la actualidad, tanto empresas como organizaciones sin ánimo de lucro e incluso los gobiernos utilizan los avances en el campo del análisis de sentimientos para saber qué es lo que piensan los usuarios de internet acerca de productos, marcas, medidas y políticas públicas, lo que convierte a esta herramienta en un instrumento valioso para medir

el respaldo social en redes sociales y otros medios electrónicos (Sobriño, 2018). En la Figura 1 se presenta la metodología aplicada para cumplir el propósito de esta investigación.

Figura 1. Procedimiento metodológico



Fuente: elaboración propia.

A continuación, se describe cada una de las etapas que aparecen en el flujo de investigación se explica de manera detallada.

Selección de palabras clave

Para recuperar los comentarios que tuvieran relación con la vacunación en tiempos de covid-19, se emplearon una serie de palabras clave (*keywords*) junto a etiquetas (*hashtags*) de Twitter. Estos términos permitieron encontrar comentarios que se relacionaban directamente con la unidad de análisis de este trabajo de investigación. En concreto, las palabras clave empleadas fueron: “vacuna”, “vacuna coronavirus” y “vacuna covid-19”.

Extracción de datos

Para extraer los comentarios de Twitter, se empleó el software estadístico R en su entorno visual Rstudio (Versión 1.2.1335); además, se emplearon las librerías de raspado web (*web scraping*) Rtweet, versión 0.7.0 (Kearney, 2016). Para raspar la información se creó una conexión con la interfaz de programación de aplicaciones (API) para programadores de Twitter. Es importante comentar que, en su versión gratuita, Twitter solamente permite extraer hasta siete días de publicaciones; por lo tanto, la extracción de comentarios se realizó semanalmente, desde el 12 de mayo hasta el 30 de septiembre de 2020, hasta lograr una muestra final de 38.034 trinos.

Procesamiento del texto

Antes de proceder a aplicar técnicas de aprendizaje de máquina para la clasificación de los comentarios en función de emociones y sentimientos, es indispensable someter los textos a una serie de técnicas que limpian y reducen las características de los mismos y facilitan su posterior análisis. Para este proceso se hizo uso del paquete informático Rapidminer 9.6.0, que ofrece varias etapas de procesamiento. En concreto, las técnicas de limpieza aplicadas fueron las siguientes: filtrado, “tokenización”, transform cases y eliminación de palabras vacías. A continuación, se explica de manera breve cada una de estas etapas. 1) filtrado: se eliminan elementos innecesarios dentro de los cuerpos de texto, como repeticiones, símbolos, signos de puntuación y espacios (Allahyari *et al.*, 2017); 2) “tokenización”: se dividen las oraciones en trozos más pequeños, lo que facilita el análisis de información mediante Machine Learning (Demidova y Klyueva, 2017); 3) *Transform cases*: todas las letras se transforman a minúsculas; 4) eliminación de palabras vacías: se eliminan de los textos palabras innecesarias que no aportan un sentido emocional; por ejemplo, conectores y artículos (Pratama *et al.*, 2019).

Aprendizaje de máquina (entrenamiento)

Una vez preparado el texto para el análisis, el paso siguiente consiste en el entrenamiento de un algoritmo de clasificación, que aprenda a reconocer emociones y sentimientos dentro de los comentarios objeto de estudio. En otras palabras, se utilizan conjuntos de datos de entrenamiento y de prueba, previamente clasificados según las emociones que presentan para enseñar a la “máquina” a identificar estos patrones en nuevos conjuntos de texto (Manning *et al.*, 2008). Para esta investigación, se empleó un conjunto de datos de entrenamiento y de prueba que ha sido ampliamente utilizado en la literatura académica: el *Wassa-2017*, en su versión en español (Mohammad y Bravo, 2017). El conjunto de datos *Wassa-2017* cuenta con 6.755 trinos previamente clasificados en cuatro sentimientos: alegría, ira, miedo y tristeza. A continuación, en la Tabla 1, se presenta la distribución del corpus.

Tabla 1. Conjunto de datos

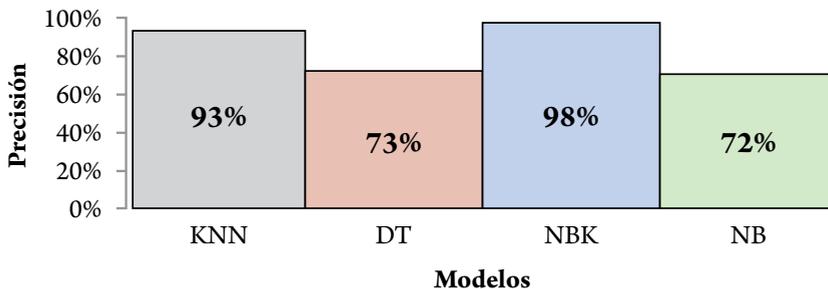
Sentimiento	Datos de entrenamiento	Datos de prueba
Alegría	823	714
Ira	857	760
Miedo	1.147	995
Tristeza	786	673

Fuente: Mohammad y Bravo (2017).

Evaluación de algoritmos (validación cruzada)

Por medio del conjunto de datos *Wassa-2017*, se construye el modelo de análisis de sentimientos utilizando el paquete informático *Rapdiminer*. No obstante, para obtener una mayor precisión, se prueban varios algoritmos de *Machine Learning* hasta obtener el mejor ajuste posible. Para esta investigación se probaron varias técnicas de clasificación integradas al software *RapidMiner*, las cuales son altamente usadas en la clasificación de sentimientos; en concreto, los algoritmos utilizados fueron los siguientes: árboles de decisión (DT), Naive Bayes (NB), Naive-Bayes Kernel (NBK) y K-Vecinos Cercanos (KNN) (Kotu y Deshpande, 2014). Cabe destacar que la evaluación de los modelos se realizó mediante validación cruzada (*cross validation*), que consiste en la partición de los comentarios en varios conjuntos de datos distintos, para estimar si la precisión es independiente de la manera en la que se dividen los datos de entrenamiento y prueba (Devi- jver y Kittler, 1982). Para este caso, los comentarios se dividieron en diez conjuntos distintos. A continuación, en la Figura 2, se presenta el desempeño de los algoritmos evaluados en esta investigación.

Figura 2. Desempeño de los algoritmos



Fuente: elaboración propia.

Como puede observarse en la Figura 2, todos los modelos obtuvieron resultados aceptables, por encima del 70%, de acuerdo con Pang y Lee (2008); sin embargo, el modelo Naive Bayes Kernel (NBK) alcanzó la mayor precisión en la clasificación, con un 98%. Lo anterior indica la idoneidad y el gran ajuste que presenta este algoritmo para los datos de entrenamiento utilizados. Vale la pena destacar que su pertinencia puede deberse a la flexibilidad que presenta, además de su característica no paramétrica, es decir, que no requiere de una distribución normal de los datos para funcionar de manera adecuada. Teniendo en cuenta lo anterior, el modelo finalmente aplicado fue el NBK.

Pesado de características

Una vez establecidos los algoritmos de aprendizaje de máquina, se identifica la aparición de los términos relacionados con las emociones de interés, como alegría, ira, miedo y tristeza (Teso *et al.*, 2018). Aunque en la literatura se presentan una gran variedad de ponderaciones, en esta investigación el que arrojó mayor precisión fue la frecuencia de términos (TF), propuesta por Luhn (1957). La TF otorga mayor relevancia a los términos con más frecuencia y evalúa el número de veces que el término aparece en el documento, tal como lo señala la siguiente ecuación, donde F_{ij} es la frecuencia del término j en el documento i :

$$P_i(t_j) = F_{ij} \quad (1)$$

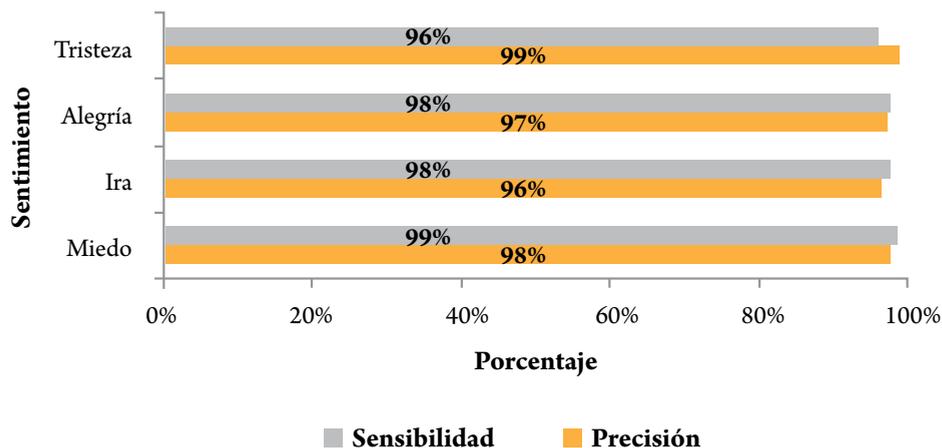
Grafos de redes sociales

Después de aplicar el análisis de sentimientos, se aplica la técnica de grafos de redes sociales con la intención de identificar en Twitter los usuarios más importantes y las emociones predominantes relacionadas con la vacunación, así como sus opiniones. Los usuarios se presentan en los gráficos como nodos y la conexión existente en ellos se identifica a través de la modularidad, lo que se refleja en distintos colores (Bastian *et al.*, 2009). Esto quiere decir que las palabras ampliamente relacionadas, o de aparición conjunta, se muestran en los gráficos del mismo color. Para adelantar el análisis de grafos se hizo uso del software Gephi 0.92.

Resultados

En los siguientes apartados se presentan los resultados obtenidos mediante los análisis aplicados a los comentarios de Twitter, empezando por la precisión y sensibilidad del algoritmo elegido (NBK), que se refleja en la Figura 3.

Figura 3. Precisión y sensibilización del modelo



Fuente: elaboración propia.

Como puede observarse en la Figura 3, la precisión y sensibilidad del modelo para clasificar correctamente nuevos datos es considerablemente alta, lo cual es ideal en un modelo de este tipo. Además, puede notarse que la precisión y la sensibilidad para cada categoría obtuvo valores que oscilan entre el 96% y el 99%. Lo anterior implica que el modelo usado presenta unos parámetros de desempeño adecuados.

Volumen de publicaciones asociadas al tema de la vacuna contra la covid-19

La cantidad de comentarios recopilados de Twitter desde el 12 de mayo hasta el 30 de septiembre del 2020 fue de 38.034. A partir de estos, se analizó que, en promedio, diariamente los usuarios realizaron 268 comentarios del tema en esta red social, con un mínimo de 10 y un máximo de 2.446 trinos diarios. Lo anterior plantea un rango y una varianza considerables en los datos. Para conocer en detalle los análisis descriptivos puede observarse la Tabla 2.

Tabla 2. Análisis descriptivos

Estadística	Valores
Media	267,8
Mediana	97,5
Desviación estándar	452,5
Varianza	204.753,6
Mínimo	10
Máximo	2.446
Suma	38.034

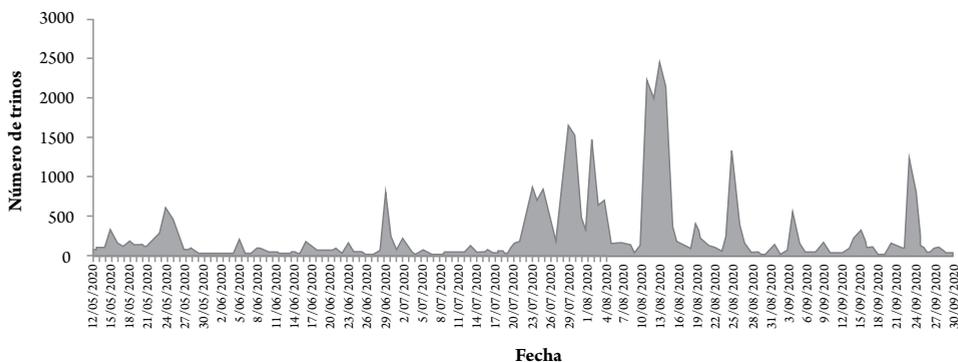
Fuente: elaboración propia a partir de los datos obtenidos de Twitter.

Es importante comentar que, frente a la veracidad de los trinos publicados, el 96% de las cuentas que participan en este debate no son verificadas y, a su vez, el 79% de los trinos publicados son difusiones de otras personas (retrinos).

Volumen de publicaciones en *El Tiempo*

En la Figura 4, se puede observar la evolución del volumen de publicaciones relacionadas con la vacunación a través del tiempo.

Figura 4. Volumen de publicaciones en Twitter



Fuente: elaboración propia a partir de los datos obtenidos de Twitter.

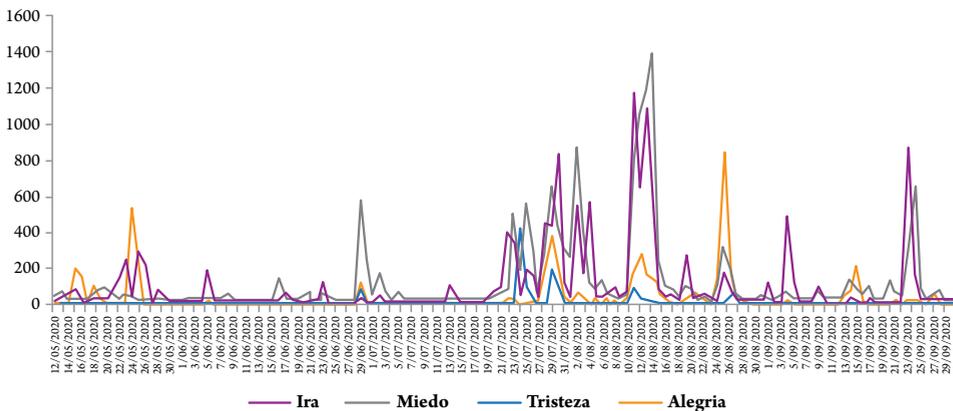
En la Figura 4, se presentan algunos elementos que no deben pasar desapercibidos. Por ejemplo, se puede ver cómo el 24 de mayo el volumen de comentarios relacionados con las vacunas aumentó de manera conside-

nable, lo que coincide con una gran cantidad de noticias difundidas ese día en medios de comunicación de alta relevancia mundial, donde se indicaba que en ese momento había más de 120 vacunas en desarrollo alrededor del planeta. Por otra parte, el mayor incremento de menciones se reporta entre el 11 y el 14 de agosto, lo que coincide con el anuncio de varios medios de comunicación frente a la aparición de una posible vacuna procedente de Rusia.

Sentimientos asociados con vacunas

A continuación, en la Figura 5, se puede observar la evolución de las emociones subyacentes en las publicaciones de Twitter analizadas a través del tiempo.

Figura 5. Evolución de las emociones en Twitter

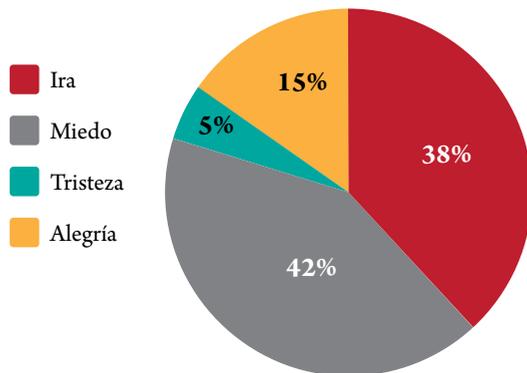


Fuente: elaboración propia a partir de los datos obtenidos de Twitter.

Puede observarse que el 24 de mayo, en el primer incremento de publicaciones, el sentimiento predominante es de alegría, ya que se asocia a la cura (vacuna) contra el virus. En esta fecha se dieron a conocer en medios nacionales e internacionales las investigaciones que se estaban realizando en búsqueda de una vacuna para el virus; sin embargo, a medida que avanzaban los días, los sentimientos más predominantes fueron el miedo y la ira. El 29 de junio el sentimiento predominante fue el miedo, en una fecha en la que medios internacionales informaron sobre experimentación china relacionada con la vacuna.

En el mes de julio y agosto, el miedo y la ira incrementaron de manera considerable, meses en los que hubo noticias relacionadas con la re-infección del virus en algunas zonas del mundo y con algunos resultados obtenidos por aquellas vacunas que estaban en experimentación. Por otro lado, para el 25 de agosto se observa que el sentimiento de alegría volvió a ser predominante, ya que las noticias enmarcadas en esta fecha daban cuenta de la llegada a fase III de prueba de algunas vacunas que estaban en desarrollo, lo cual indica su éxito en las fases anteriores. Las fechas del 24 al 25 de septiembre aparecieron de nuevo los sentimientos de miedo e ira como predominantes, lo que puede asociarse con algunas noticias donde se comentaban resultados fallidos para las vacunas en desarrollo. A continuación, en la Figura 6, se observa la distribución de emociones a lo largo del tiempo.

Figura 6. Distribución de las emociones en Twitter



Fuente: elaboración propia a partir de los datos obtenidos de Twitter.

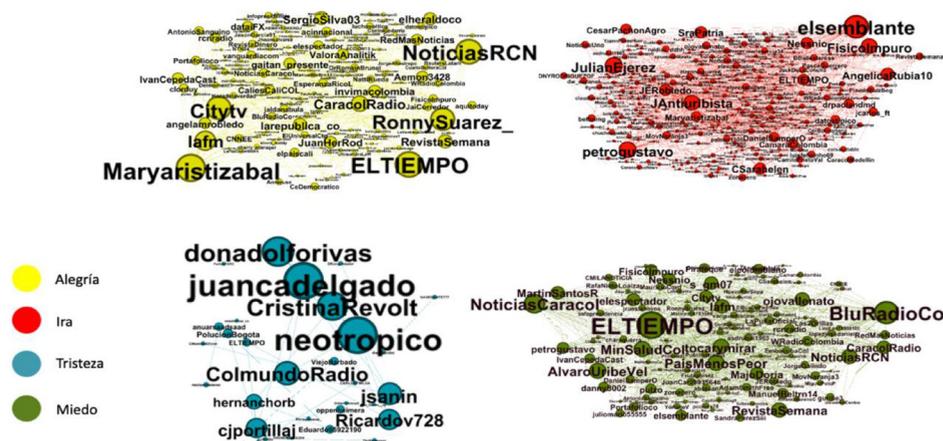
Con un 42%, el sentimiento predominante en los usuarios de Twitter con respecto a la vacuna fue el miedo. Así, del total de comentarios analizados, 16.086 presentan al miedo como emoción principal; en segundo lugar aparece la ira, con el 38% (14.428 comentarios); en tercer lugar, la alegría, con una proporción del 15% (5.769 comentarios), y, finalmente, la tristeza, con el menor porcentaje, de tan solo un 5% (1.751 comentarios).

Actores relevantes

Por medio del paquete informático Gephi, se determinan en la Figura 7 los actores más relevantes en Twitter respecto de la vacunación contra la

covid-19. Esto se logra a partir del volumen de interacciones que reciben estas cuentas por parte de sus seguidores, como retrinos, “me gusta” (*like*) y comentarios. Posteriormente, mediante las técnicas de análisis de sentimientos aplicadas en los apartados anteriores, se identifican y se clasifican los más importantes que estos usuarios mostraron ante la vacunación en esta red social. Cabe destacar que en las cuentas analizadas pueden aparecer varios sentimientos al mismo tiempo; no obstante, el tamaño del nodo en cada grupo indica la importancia que tiene esta cuenta para cada grupo de sentimientos. Una cuenta puede aparecer en dos o más sentimientos a la vez, ya que sus trinos pueden aportar a diferentes sentimientos y ser relevantes para cada una de estas categorías.

Figura 7. Actores relevantes



Fuente: elaboración propia a partir de los datos obtenidos de Twitter.

De acuerdo con la Figura 7, las comunidades de usuarios detectadas en los datos presentaban informantes claros que difundían opiniones sobre la vacuna covid-19, en cuentas de Twitter, en su mayoría, adscritas a medios de comunicación, como radio, prensa y televisión, así como a algunos influenciadores y políticos, principalmente congresistas. En el grupo de ira (en color rojo), se distinguen varias cuentas de la oposición al gobierno colombiano, entre las más importantes y reconocidas, la cuenta @petrogustavo del senador Gustavo Petro, líder actual de la oposición

al gobierno; y las cuentas @fisicoimpuro, @elsemblante y @julianejerez, de líderes del progresismo opositor. En cuanto a periodistas, una de las cuentas más visibles en este grupo fue la de @danielsamperO, reconocido comunicador e influenciador crítico del gobierno actual. Tras una revisión cualitativa de algunos trinos difundidos por los integrantes de estas cuentas, pudo notarse que el sentimiento de ira correspondía a las fuertes críticas al gobierno nacional que realizaron por cuenta del manejo de la pandemia. También critican las negociaciones internacionales para la consecución de la vacuna para el país. En el otro extremo, y profesando sentimientos mayoritarios de alegría, se encuentran las cuentas oficiales de los medios de comunicación, ampliamente reconocidos en el país, como @el-Tiempo, @NoticiasRCN y @CityTV. También figuran periodistas de estos mismos medios, como @Maryaristizabal y @RonnySuarez. La revisión cualitativa de algunos trinos en este grupo permitió identificar que estos actores se relacionaban con la alegría en la medida que se mostraban positivos frente a los avances en el desarrollo de vacunas covid-19. Frente a la emoción de tristeza, resalta que los principales exponentes de este sentimiento eran los influenciadores y no medios de comunicación ni políticos reconocidos. Se trataba de cuentas con un alto número de seguidores que lograron transmitir sentimientos de tristeza asociados con la vacunación, debido al panorama de dificultades que vivía actualmente la sociedad colombiana producto de la covid-19.

Finalmente, y como máximos exponentes del miedo, se observan nuevamente medios de comunicación reconocidos, como @ElTiempo, @BlueRadio, @NoticiasCaracol y @RevistaSemana, y, sorpresivamente, también aparecen cuentas oficiales del gobierno nacional, como @MinSalud y el expresidente Álvaro Uribe, con su cuenta @AlvaroUribeVel, un personaje público muy cercano al gobierno actual.

A continuación, en la Tabla 3 se evidencian algunos ejemplos de trinos difundidos por los actores analizados que contribuyeron a los diferentes sentimientos identificados.

Tabla 3. Análisis descriptivos

Actor	Trino	Sentimiento principal
@PetroGustavo	“Por que en Cuba se elabora una vacuna y por qué en Colombia no? Porque a quienes han dirigido el estado colombiano jamás les ha interesado el desarrollo del saber en la sociedad colombiana.”	Ira
@NoticiasRCN	“Una de las noticias más importantes del día es que la Universidad de Oxford retomó ensayos de la vacuna contra el covid-19, así van los avances.”	Alegría
@jsanin	“La vacuna va a encontrar un país sin rastros de salud mental por el encierro.”	Tristeza
@ElTiempo	“Las pruebas de la vacuna que desarrolla la Universidad de Oxford con el laboratorio AstraZeneca se suspendieron debido a una grave reacción adversa en uno de los pacientes que hacen parte del grupo en el que se realizan los estudios.”	Miedo

Fuente: elaboración propia.

Discusión y conclusiones

Del análisis de sentimientos y de grafos llevado a cabo se pueden extraer conclusiones importantes para comprender los comportamientos y sentimientos de usuarios de Twitter hacia la vacuna de la covid-19 e identificar algunos aspectos metodológicos que vale la pena ser resaltados con el ánimo de enriquecer futuras investigaciones de este tipo. Para empezar, y respecto de los aspectos metodológicos, se puede concluir que la base de datos empleada para entrenar y probar el algoritmo de aprendizaje de máquina (Wassa, 2017) permite un ajuste excelente al modelo de clasificación de un cuerpo de datos en español. En términos generales, todos los modelos revisados ofrecieron niveles de exactitud por encima del 70%, siendo el Naive Bayes Kernel el que arrojó un ajuste más alto, con un nivel del 98% de precisión. Esto puede deberse a la particularidad de este algoritmo y, específicamente, a su condición no paramétrica. Es pertinente mencionar que, si bien en el estudio de Arango y Osorio (2021) se prueban diferentes algoritmos para evaluar la precisión del conjunto de datos Wassa-2017 a un corpus en español, hasta el momento esta investigación es la primera en aplicar un modelo de validación cruzada a este cuerpo de datos en español. Por otra parte, en la investigación anteriormente mencionada el mejor

algoritmo identificado fue el Deep Learning, con un nivel de precisión del 95,8%, lo que implica que el Naive Bayes Kernel mediante validación cruzada ofrece mejores resultados.

Respecto a los comentarios vertidos en Twitter relacionados con la vacuna contra la covid-19, pudo observarse que el volumen de estos aumentaba de manera simultánea con las noticias internacionales relacionadas con el estado de desarrollo de las vacunas. Al analizar los picos de tendencia, se concluye que estos se correspondían con grandes despliegues noticiosos en medios de comunicación internacionales relacionados con este tema. Además, los sentimientos que se asociaban a esos comentarios también coincidían con noticias que lucían esperanzadoras o traían miedo e incertidumbre. Para ejemplificar lo anteriormente expuesto, se puede apreciar en las gráficas que, para la fecha del 24 de mayo de 2020, el sentimiento predominante era la alegría, lo que corresponde con diferentes noticias difundidas en medios internacionales, como CNN (2020), donde indicaban que, para ese entonces, más de 120 vacunas contra el coronavirus se encontraban en desarrollo. Por otra parte, para finales de julio e inicios de agosto de 2020 se observa que emociones como la ira y el miedo eran las más recurrentes en las opiniones de Twitter, aspecto que se relaciona con una gran cantidad de noticias difundidas por diferentes medios donde se informaba acerca de una posible segunda ola de infección por covid-19 (El País, 2020; The New York Times, 2020b).

De manera general, la mayoría de los comentarios asociados con la vacuna presentaron sentimientos negativos, lo que necesariamente no quiere decir que la mayoría de las personas rechazaran la vacunación como una medida de protección ante el virus, sino que la incertidumbre y el miedo, producto del escenario desesperanzador de pandemia, eran más fuertes que el optimismo. Además, para la fecha en que se recolectaron los datos de este estudio, aún no existían vacunas contra la covid-19 ampliamente probadas y exitosas. Otro factor que, sin lugar a duda, puede estar relacionado con la prevalencia de sentimientos negativos como el miedo, en un escenario donde se esperaría una amplia aceptación de la vacuna, puede estar ligado a la alta tasa de noticias falsas difundidas en diversos medios de comunica-

ción. Autores como Montagni *et al.* (2021) sugieren que durante la pandemia por covid-19 la difusión de noticias falsas se convirtió en una verdadera epidemia de comunicación, que se calificó de “infomedia”, lo que explicaría el peso de sentimientos negativos, principalmente de miedo.

Del análisis de los usuarios relevantes en Twitter respecto del tema de vacunación, es importante mencionar que, en su mayoría, estas cuentas pertenecían a usuarios verificados como medios de comunicación, políticos e influenciadores reconocidos, lo que pone en evidencia la importancia y responsabilidad de estos actores mediáticos en cuanto a la aceptación o rechazo de un tema tan sensible como la vacuna contra la covid-19 por parte de la comunidad en general. De acuerdo con los resultados, se puede observar que la oposición cumple su rol de crítica al gobierno, plasmando en sus trinos con sentimientos asociados más que todo con la ira, al mismo tiempo que los medios de comunicación reconocidos en Colombia aportaban en sus intervenciones por esta red social sentimientos asociados con la alegría, por ejemplo, al celebrar el éxito de la ciencia con el desarrollo de vacunas para prevenir y mitigar los efectos del virus. No obstante, algo que sorprende un poco en este estudio es la participación de ciertos medios de comunicación oficiales, así como de cuentas verificadas del gobierno nacional en la difusión de trinos que se asocian con el sentimiento del miedo. Estos resultados podrían emplearse para tomar con más cautela la información que difunden este tipo de actores mediáticos, ya que contribuye a generar pánico colectivo.

Limitaciones y futuras líneas de investigación

En esta investigación se analizan trinos de aproximadamente cuatro meses (desde el 12 de mayo hasta el 30 de septiembre de 2020), antes de que las vacunas alcanzaran su fase de comercialización y, por lo tanto, empezaran a ser probadas de manera masiva en humanos. De esta manera, los resultados del análisis de sentimientos pueden variar de manera considerable ahora que las vacunas han demostrado su efectividad y que los gobiernos alrededor del mundo implementaron las primeras fases de inmunización de su población. Sería interesante que futuros estudios pudiesen adelantar investigaciones de carácter longitudinal para contrastar los resultados obtenidos

aquí con periodos más recientes donde la dinámica ha cambiado de manera considerable. Por otro lado, también se podrían adelantar estudios multiculturales para contrastar las emociones difundidas por usuarios de la red social Twitter en otros países y, de esta forma, verificar si existen diferencias socioculturales en la respuesta al tema de la vacuna. Lo anterior podría ayudar a identificar estrategias mediáticas para generar mayor tranquilidad y respaldo de los usuarios a medidas tan importantes como la vacunación.

Una limitación adicional que presenta este estudio tiene que ver con la red social escogida para adelantar el análisis de información. El público objetivo que atienden las redes sociales varía de manera significativa entre sí; por lo tanto, limitarse a una sola red social como Twitter puede representar un sesgo importante respecto a la población analizada. Futuras investigaciones podrían aplicar análisis similares en otras plataformas de opinión, como Facebook, YouTube e incluso la recientemente popular Tik-Tok. Esto podría contribuir a encontrar diferencias en la manera de percibir el tema de la vacunación por parte de distintos perfiles de usuarios.

Referencias

- Allahyari, M. *et al.* (2017). A brief survey of text mining: Classification, clustering and extraction techniques. *ArXiv*, 1, 1-13. <http://arxiv.org/abs/1707.02919>
- Aps, L. *et al.* (2018). Eventos adversos de vacinas e as consequências da não vacinação: uma análise crítica. *Revista de Saude Pública*, 52, 40. <https://doi.org/10.11606/S1518-8787.2018052000384>
- Arango Pastrana, C. A. y Osorio Andrade, C. F. (2021). Aislamiento social obligatorio: un análisis de sentimientos mediante *machine learning*. *Suma de Negocios*, 12(26), 1-13. <https://doi.org/10.14349/sumneg/2021.v12.n26.a1>
- Ashton, J. (2021). Covid-19 and the anti-vaxxers. *Journal of the Royal Society of Medicine*, 114(1), 42-43. <https://doi.org/10.1177/0141076820986065>

- Atalan, A. (2020). Is the lockdown important to prevent the Covid-9 pandemic? Effects on psychology, environment and economy-perspective. *Annals of Medicine and Surgery*, 56, 38-42. <https://doi.org/10.1016/j.amsu.2020.06.010>
- Bastian, M., Heymann, S. y Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks [Conference presentation]. Third International AAAI Conference on Weblogs and Social Media, 361-362. <https://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>
- Carrasco, I. R. Z. y Lozano, J. C. (2018). Grupos antivacunas: el regreso global de las enfermedades prevenibles. *Revista Latinoamericana de Infectología Pediátrica*, 31(1), 17-21. <https://www.medigraphic.com/pdfs/infectologia/lip-2018/lip181d.pdf>
- Carrieri, V., Madio, L. y Principe, F. (2019). Vaccine hesitancy and (fake) news: Quasi-experimental evidence from Italy. *Health Economics*, 28(11), 1377-1382. <https://doi.org/10.1002/hec.3937>
- CNN (2020). ¿En qué va la vacuna contra el coronavirus? 8 de junio. <https://cnnespanol.cnn.com/2020/06/08/en-que-va-la-vacuna-contra-el-coronavirus/>
- Crokidakis, N. (2020). Covid-19 spreading in Rio de Janeiro, Brazil: do the policies of social isolation really work? *MedRxiv*, 136, 109930. <https://doi.org/10.1016/j.chaos.2020.109930>
- Demidova, L. y Klyueva, I. (2017). Improving the classification quality of the SVM classifier for the imbalanced datasets on the base of ideas the SMOTE algorithm. *ITM Web of Conferences*, 10, 02002. <https://doi.org/10.1051/itmconf/20171002002>
- Devijver, P. A. y Kittler, J. (1982). *Pattern recognition: A statistical approach*. Prentice Hall International.

- Dubey, A. (2021). Public sentiment analysis of Covid-19 vaccination drive in India. Jaipuria Institute of Management. <https://doi.org/10.2139/ssrn.3772401>
- El País (2020). ¿Qué dicen los datos de agosto sobre la nueva expansión del coronavirus en España? 23 de agosto. <https://elpais.com/sociedad/2020-08-22/que-dicen-los-datos-de-agosto-sobre-la-nueva-expansion-del-coronavirus-en-espana.html>
- European Centre for Disease Prevention and Control (2020,). Systematic scoping review on social media monitoring methods and interventions relating to vaccine hesitancy. ECDC, 9 de marzo. <https://www.ecdc.europa.eu/en/publications-data/systematic-scoping-review-social-media-monitoring-methods-and-interventions>
- Fernández, J. y Baquero, H. (2019). El movimiento anti-vacunas y la anti-ciencia como amenaza para la salud pública. *Salud*, 51(2), 104-107. <https://doi.org/10.18273/revsal.v51n2-2019002>
- Hussein, E., Prerna Juneja, P. y Mitra, T. (2020). Measuring misinformation in video search platforms: An audit study on YouTube. [Conference] The 23rd ACM Conference on Computer-Supported Cooperative Work and Social Computing - CSCW '20At: Minneapolis, MN, USA. <https://doi.org/10.1145/3392854>
- Hotez, P. (2019). America and Europe's new normal: the return of vaccine-preventable diseases. *Pediatric Research*, 85(7), 912-914. <https://doi.org/10.1038/s41390-019-0354-3>
- Jamison, A. *et al.* (2020). Adapting and extending a typology to identify vaccine misinformation on Twitter. *American Journal of Public Health*, 110(S3), S331-S339. <https://doi.org/10.2105/AJPH.2020.305940>
- Kearney, M. W. (2016). Package 'rtweet'. <https://cran.r-project.org/web/packages/rtweet/rtweet.pdf>

- Kotu, V. y Deshpande, B. (2014). *Predictive analytics and data mining: concepts and practice with rapidminer*. Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-801460-8.00013-6>
- Kuckertz, A. et al. (2020). Startups in times of crisis – A rapid response to the Covid-19 pandemic. *Journal of Business Venturing Insights*, 13, e00169. <https://doi.org/10.1016/j.jbvi.2020.e00169>
- Liu, B. (2010). Sentiment analysis and subjectivity. *Handbook of Natural Language Processing*, 2, 627-666.
- Luhn, H. P. (1957). A statistical approach to mechanized encoding and searching of literary information. *IBM Journal of Research and Development*, 1(4), 309-317. <https://doi.org/10.1147/rd.14.0309>
- Manning, C., Raghavan, P. y Schütze, H. (2008). Chapter 13: Text classification and naive bayes. *Introduction to Information Retrieval*, 234-265. <https://doi.org/10.1017/CBO9780511809071.014>
- Mohammad, S. y Bravo, F. (2017). WASSA-2017 shared task on emotion intensity. 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (pp. 34-49). Association for Computational Linguistics. <https://doi.org/10.18653/v1/W17-5205>
- Montagni, I. et al. (2021). Acceptance of a Covid-19 vaccine is associated with ability to detect fake news and health literacy. *Journal of Public Health*, 43(4), 695-702. <https://doi.org/10.1093/pubmed/fdab028>
- Mukherjee, K. (2020). Covid-19 and lockdown: Insights from Mumbai. *Indian Journal of Public Health*, 64, S168-S171. https://doi.org/10.4103/ijph.IJPH_508_20
- Murphy, J. et al. (2021). Psychological characteristics associated with Covid-19 vaccine hesitancy and resistance in Ireland and the Uni-

ted Kingdom. *Nature Communications*, 12(1), 1-16. <https://doi.org/10.1038/s41467-020-20226-9>

Musinguzi, G. y Asamoah, B. O. (2020). The science of social distancing and total lock down: Does it work? Whom does it benefit? *Electronic Journal of General Medicine*, 17(6), 17-19. <https://doi.org/10.29333/ejgm/7895>

Neumann-Böhme, S. *et al.* (2020). Once we have it, will we use it? A European survey on willingness to be vaccinated against Covid-19. *European Journal of Health Economics*, 21(7), 977-982. <https://doi.org/10.1007/s10198-020-01208-6>

OMS (2020). Draft landscape of Covid-19 candidate vaccines. 20 April 2020. <https://www.who.int/blueprint/priority-diseases/key-action/novel-coronavirus-landscape-ncov.pdf>

Ortiz, E. (2020). Tres de cada cuatro mexicanos aceptarían aplicación de vacuna contra covid-19. UDG TV, 1 de septiembre. <http://udgtv.com/noticias/tres-de-cada-cuatro-mexicanos-acceptarian-aplicacion-de-vacuna-contra-covid-19/>

Pang, B. y Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135. <https://doi.org/10.1561/1500000011>

Paul, E., Steptoe, A. y Fancourt, D. (2021). Attitudes towards vaccines and intention to vaccinate against Covid-19: Implications for public health communications. *The Lancet Regional Health-Europe*, 1, 100012. <https://doi.org/10.1016/j.lanepe.2020.100012>

Pratama, B. T., Utami, E. y Sunyoto, A. (2019). The impact of using domain specific features on lexicon based sentiment analysis on Indonesian app review. [International Conference] Information and Communications Technology (Icoiact), 24-25 de julio. <https://doi.org/10.1109/ICOIACT46704.2019.8938419>

- Piedrahíta-Valdés, H. *et al.* (2021). Vaccine hesitancy on social media: Sentiment analysis from June 2011 to April 2019. *Vaccines*, 9(1), 28. <https://doi.org/10.3390/vaccines9010028>
- Salleras, L. (2018). Movimientos antivacunas: una llamada a la acción. *Vacunas*, 19(1), 1-3. <https://doi.org/10.1016/j.vacun.2018.03.001v>
- Sobrino, J. C. S. (2018). *Análisis de sentimientos en Twitter. Inteligencia artificial*. Universidad Oberta de Catalunya.
- Suárez, V. M. (2020). Vacuna frente a la covid-19: administración imperativa vs. información convincente. *Pediatría Integral*, 24(8), 431-434.
- Sued, G. (2020). El algoritmo de YouTube y la desinformación sobre vacunas durante la pandemia de covid-19. *Chasqui*, 1(145), 163-180. <https://doi.org/10.16921/chasqui.v1i145.4335>
- Teso, E., Olmedilla, M., Martínez-Torres, M. R. y Toral, S. L. (2018). Application of text mining techniques to the analysis of discourse in eWOM communications from a gender perspective. *Technological Forecasting and Social Change*, 129, 131-142. <https://doi.org/10.1016/j.techfore.2017.12.018>
- The New York Times (2020a). Coronavirus vaccine tracker. <https://www.nytimes.com/interactive/2020/science/coronavirus-vaccine-tracker.html>
- The New York Times (2020b). “Allá vamos otra vez”: España vive una segunda ola de coronavirus. 31 de agosto. <https://www.nytimes.com/es/2020/08/31/espanol/mundo/rebote-espana.html>
- Trogen, B., Oshinsky, D. y Caplan, A. (2020). Adverse consequences of rushing a Sars-CoV-2 vaccine: Implications for public trust. *JAMA*, 323(24), 2460-2461. <https://doi.org/10.1001/jama.2020.8917>

Wolfe, R. M. (2002). Anti-vaccinationists past and present. *BMJ*, 325(7361), 430-432. <https://doi.org/10.1136/bmj.325.7361.430>

Worldometer. (2020). Countries where Covid-19 has spread. <https://www.worldometers.info/coronavirus/countries-where-coronavirus-has-spread/>

Zucker, J. R. *et al.* (2020). Consequences of undervaccination—measles outbreak, New York City, 2018-2019. *New England Journal of Medicine*, 382(11), 1009-1017. <https://doi.org/10.1056/NEJMoa1912514>