

MINERIA DE DATOS: ANÁLISIS DE SENTIMIENTO EN TWITTER BASADO EN LEXICONES SOBRE EL USO DE DIÓXIDO DE CLORO PARA EL TRATAMIENTO DEL COVID-19

Enrique Edgardo Condor Tinoco

Universidad Nacional José María Arguedas
Andahuaylas, Apurímac, Perú
<https://orcid.org/0000-0002-2711-9517>

José Antonio Rojas Cusi

Universidad Nacional José María Arguedas
Andahuaylas, Apurímac, Perú

Anthony Zevallos Rodríguez

Universidad Nacional José María Arguedas
Andahuaylas, Apurímac, Perú

Carlos Yinnel Castro Buleje

Universidad Nacional José María Arguedas
Andahuaylas, Apurímac, Perú
<https://orcid.org/0000-0002-9368-3634>

All content in this magazine is licensed under a Creative Commons Attribution License. Attribution-Non-Commercial-Non-Derivatives 4.0 International (CC BY-NC-ND 4.0).



Resumen: COVID-19 según la Organización Mundial de Salud OMS es la “enfermedad infecciosa causada por el coronavirus que se ha descubierto más recientemente, para el tratamiento del COVID-19 se han informado por diferentes medios diversos productos, entre ellas se tiene al dióxido de cloro, se dice que, puede curar y prevenir el COVID-19 porque incrementa los niveles de oxígeno en la sangre; también expertos en la salud indicaron que es falso y, por el contrario, esta puede oxidar la hemoglobina y desencadenar más problemas de salud. Se planteó el objetivo de desarrollar un modelo de minería de datos sobre el Análisis de Sentimiento en Twitter basado en lexicones sobre el uso de dióxido de cloro para el tratamiento del COVID-19, para su desarrollo se usó la metodología CRISP-DM, y la herramienta Orange Canvas. En los resultados se obtuvo que la palabra dióxido de cloro está asociada a las palabras, si, tomar y Covid; por otro lado, la opinión de las personas en relación al dióxido de cloro con tendencia a negatividad y neutralidad, dan como resultado calificaciones compuestas, lo que significa que 4500 Tweets de los usuarios de Twitter tienen opiniones de neutralidad y rechazo en distintas proporciones sobre el tratamiento del dióxido de cloro para el COVID-19.

Palabras clave: *análisis de sentimiento, dióxido de cloro, Twitter, COVID-19.*

INTRODUCCIÓN

Twitter fue fundado por Jack Dorsey, es un servicio de microblogueo y de red social, que se encuentra continuamente alimentado por millones de personas desde diferentes partes del mundo. Es una plataforma que limita sus mensajes a 280 caracteres. Twitter es una plataforma en línea para publicar y leer mensajes de texto de hasta 140 caracteres (Orihuela, 2011). A diferencia de las redes sociales de Facebook y LinkedIn las relaciones

entre los extremos de los usuarios no se ponen en contacto directo, en el caso de twitter se tiene a seguidores (follower) y los seguidos (followed), de esta manera como que se tiene 2 listas en Twitter, las lista de los que te siguen y la lista de a quienes sigues; Si sigues a alguien, eso quiere decir que verás sus “tweets” (mensajes que escribe en Twitter) en tu “cronología” (o “timeline”, conjunto ordenado en el tiempo de los mensajes que te llegan de la gente a quienes sigues) (Rubín, 2020).

La presente investigación trata del tema del uso del dióxido de cloro para el tratamiento del COVID-19. El COVID-19 según la Organización Mundial de Salud OMS es la “enfermedad infecciosa causada por el coronavirus que se ha descubierto más recientemente; este nuevo virus como la enfermedad que provoca eran desconocidos antes de que estallara el brote en Wuhan (China) en diciembre de 2019” (Organización Mundial de Salud, 2020); asimismo, la OMS ya lo declaró como una pandemia que afecta a muchos países de todo el mundo. Para el tratamiento del COVID-19, por la información que han publicado en diferentes medios, hay gente que está consumiendo el dióxido de cloro para tratarse del COVID-19; el dióxido de cloro es un gas de color amarillo-rojizo que se descompone muy rápido el aire, se usa como blanqueador en las fábricas que producen papel y en las plantas de tratamiento de agua que producen agua potable, el dióxido de cloro forman iones de clorito, son sustancias muy reactivas, por lo tanto, le permite matar a bacterias y otros microorganismos en el agua (Agencia para Sustancias Tóxicas y el Registro de Enfermedades, 2004).

Se han advertido en diferentes medios sobre el uso del dióxido de cloro ya que se dice que, puede curar y prevenir el COVID-19 porque incrementa los niveles de oxígeno en la sangre; también expertos en la salud indicaron que es falso y, por el

contrario, esta puede oxidar la hemoglobina y desencadenar más problemas de salud. En la página web de As Perú público que, el Ministerio de Salud del Perú MINSA prohíbe la venta de dióxido de cloro para tratar el COVID-19 señalando que “El consumo de dióxido de cloro o clorito de sodio puede causar graves daños a la salud de las personas, pues no se trata de un medicamento sino de un poderoso agente blanqueador, por lo tanto, no cura ninguna enfermedad” (AS Perú, 2020). En la web de la BBC NEWS se publicó que, este producto conocido como “Suplemento Mineral Milagroso” lleva ya varios años publicitándose como un remedio para muchas enfermedades incluso el cáncer; Sin embargo, ninguna institución sanitaria lo reconoce como medicamento y en la actualidad con la pandemia del COVID-19 que azota al mundo, vuelva a aparecer como algo capaz de prevenir los efectos del virus (Orgaz, 2020). En la página web de la CNN español se publicó que, la Cámara de Diputados de Bolivia aprobó un proyecto de ley que regula de manera excepcional la elaboración, comercialización, y uso consentido del dióxido de cloro, esto de manera preventiva y como tratamiento para pacientes con COVID-19 (Trucco, 2020). En la investigación denominada “Eficacia y seguridad del dióxido de cloro para el tratamiento de COVID-19”, concluyeron los autores en: a) En la búsqueda de la literatura, no encontraron ningún estudio que muestre la seguridad y eficacia de esta sustancia para el tratamiento del COVID-19, b) Diversas agencias gubernamentales de nuestra región (Chile, Argentina, Uruguay, Paraguay y Ecuador), recomiendan no consumir esta sustancia para el tratamiento de COVID-19, debido a que no se ha demostrado ningún beneficio terapéutico, por el contrario, ha mostrado eventos adversos graves (Cabezas Sanchez, Solari Zerpa, & Caballero Ñopo, 2020).

El tema del uso del dióxido de cloro para el tratamiento del COVID-19 se ha difundido de manera masiva en Twitter y las opiniones son diversas, justamente por las diferentes noticias que se ha dicho de este producto tal como se detalla en el párrafo anterior. Con las herramientas que posee el software Orange Canvas 3.26 y los datos que se encuentran en Twitter sobre el uso del dióxido de cloro se puede realizar minería de datos específicamente en el campo del análisis de sentimiento. La minería de datos (MD) es un intento de buscarle sentido a la explosión de información que actualmente puede ser almacenada (Mitra & Acharya, 2005).

En función a todo el análisis que se ha realizado en los párrafos anteriores, se plantea el objetivo de desarrollar un modelo de minería de datos sobre el análisis de sentimiento basado en lexicones sobre el uso del dióxido de cloro para el tratamiento de COVID-19, haciendo uso de la metodología CRISP-DM, modelo estándar abierto del proceso, en inglés “Cross Industry Standard Process for Data Mining” que divide el proceso de minería de datos en seis fases principales, los que son explicados en la parte de metodología de la presente investigación.

Para cumplir con el objetivo planteado y obtener los resultados, en la introducción se ha descrito el problema, los objetivos y las herramientas y metodologías que se usaran para desarrollar la investigación. En la parte de la metodología, siguiendo los pasos de CRISP-DM se desarrolla el modelo de minería de datos sobre el análisis de sentimiento en Twitter basado en lexicones sobre el uso del dióxido de cloro para el tratamiento de COVID-19; después se presentan los resultados obtenidos, para luego finalizar con las conclusiones de la investigación.

METODOLOGÍA

El desarrollo de la investigación se centra en los pasos de la metodología CRISP-DM (comprensión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación y despliegue), la secuencia se sigue según lo descrito en la Figura 1, la herramienta usada para el procesamiento de los datos fue Orange Canvas, los datos fueron tomados de la red social Twitter a través del widget Twitter que el software posee, el tema es sobre el análisis del sentimiento basado en el lexicones sobre el dióxido de cloro como posible tratamiento para el COVID-19.



Figura 1. Pasos de la Metodología CRISP-DM.

Fuente: (Vallalta Rueda, s.f.).

a) Comprensión del negocio.

Se comenta en diferentes medios sobre el uso del dióxido de cloro ya que se dice que, puede curar y prevenir el COVID-19 porque incrementa los niveles de oxígeno en la sangre; también expertos en la salud indicaron que es falso. El tema del uso del dióxido de cloro para el tratamiento del COVID-19 se ha difundido de manera masiva en Twitter y las opiniones son diversas, justamente es por las diferentes noticias que se ha dicho sobre este producto.

Es preciso analizar el sentimiento de los comentarios realizados e identificar lo que piensa la población sobre este producto.

b) Compresión de los datos.

• Widgets usados:

Los widgets de Orange Canvas usados en el tratamiento de los datos fueron:

Twitter: permite consultar tweets a través de la API de Twitter. Puede consultar por contenido y autor o ambos y acumular resultados.



Twitter

Data Table: recibe uno o más conjuntos de datos en su entrada y los presenta como una hoja de cálculo.



Data Table

Preprocess Text: divide el texto en unidades más pequeñas (tokens), las filtra, ejecuta la normalización (derivación, lematización).



Preprocess Text

Word cloud: muestra tokens en el corpus, su tamaño indica la frecuencia de la palabra en el corpus o el recuento promedio de una bolsa de palabras.



Word Cloud

Sentiment Analysis: predice el sentimiento para cada documento en un corpus. Utiliza los módulos de sentimiento de Liu Hu y Vader de NLTK.



Sentiment Analysis

Corpus viewer: está diseñado para ver archivos de texto, siempre generará una instancia de corpus.



Corpus Viewer

Select Columns: el usuario decide qué atributos se utilizarán. Orange distingue entre atributos ordinarios, atributos de clase (opcionales) y meta atributos.



Select Columns

FreeViz: ayuda a visualizar los resultados: los puntos de la misma clase se atraen entre sí, los de diferente clase se repelen



FreeViz

• Extracción de los datos

Los datos fueron recopilados directamente de Twitter haciendo el uso de su API que está incorporado en el mismo widget Twitter, lo cual una vez obtenida los Tweets se seleccionó las columnas importantes para nuestro análisis, quedando las siguientes columnas con las cuales se trabajó: autor, Tweet, descripción del autor y fecha y hora. Para empezar la consulta se debe tener las credenciales del API de Twitter, de ser el caso que no cuente con estos datos, se debe hacer una solicitud de acceso como desarrollador desde la internet. Para la selección de los tweets se usó la palabra clave “Dióxido de cloro” con el fin de obtener mejores resultados en tiempo real ya que el widget Twitter nos permite realizar este proceso, se usó este criterio para obtener consultas más precisas, ya que está palabra se encuentra en tendencia en estos tiempos de pandemia debido a su polémica como tratamiento para el COVID-19; se estableció analizar 4500 Tweets.

c) Preparación de los datos.

En la extracción de los datos, se observa que hay caracteres extraños, los cuales no servirían para el procesamiento y análisis. Para la limpieza de los datos se usa el widget Preprocess Text con lo cual se remueve, urls, acentos, códigos htmls, convertir a minúsculas, se tokeniza, etc. así como también se restringirá los caracteres extraños, con el objetivo de garantizar el trabajo con datos limpios. En un bloc de notas se lista los caracteres o palabras que se desea excluir y deben ser cargados al Preprocess Text como se observa en Figura 3.

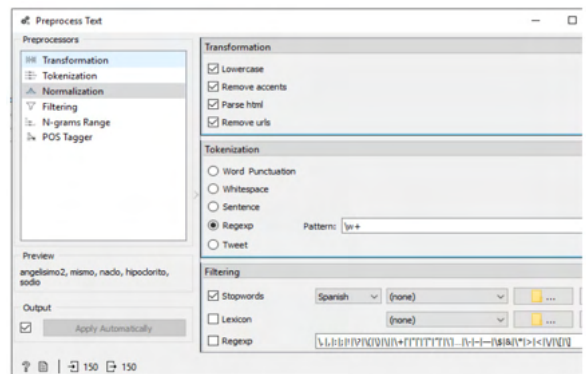


Figura 3. Interface de inicio del Preprocess Text.

Fuente: elaboración propia en Orange Canvas.

Variables	Author	Content	Date	Language	Location
<input checked="" type="checkbox"/> Show variable labels (if present)	1 @algotagniaaa	RT @IdeasCongresoCL: @floricmatotuda El clorito de sodio, al activarse con un ácido ligero, se transforma en dióxido de clor...	2020-10-11 ...	es	?
<input type="checkbox"/> Visualize numeric values	2 @NOVOLVNU...	@javierlanari TODAVIA ESTAMOS ESPERANDO SOLIDARIDAD CON LA FAMILIA DEL NIÑO MUERTO POR TOMAR EL DIOXIDO...	2020-10-11 ...	es	?
<input checked="" type="checkbox"/> Color by instance classes	3 @EfectoCocuyo	#CocuyoCheques No hay ninguna evidencia médico-científica que demuestre que tomar dióxido de cloro puede curar ...	2020-10-11 ...	es	?
Selection	4 @mihidalgo	RT @alfrepared: Las películas norteamericanas de Netflix son un cóctel de inclusión y corrección política tan fuertes que sab...	2020-10-11 ...	es	?
<input checked="" type="checkbox"/> Select full rows	5 @CariucCarina	@javierlanari TAMBIEN FALLARON LOS MEDIOS POR SU FALTA DE COMPROMISO PORQUE UNA PERSONA AL AIRE TOMO ...	2020-10-11 ...	es	?
	6 @wada_3M	@Angelisimo2 Allí no va Killgates, además saben utilizar el Dióxido de Cloro. (Malaria)	2020-10-11 ...	es	?
	7 @ymyigue	No aguanto más venga el dióxido de cloro https://t.co/ePKI25Wv	2020-10-11 ...	es	?
	8 @introdagro	@HenyChinaski67 Que tomen dióxido de cloro para aclarar la mente	2020-10-11 ...	es	?
	9 @ESAArgentina	Coronavirus en Argentina: diputada de J.C recomendó usar dióxido de cloro y generó fuerte repudio Cámara de Diputados, ...	2020-10-11 ...	es	?
	10 @calaveritabike	@pasotti_ En el mio hablan de las maravillas de tomar dióxido de cloro	2020-10-11 ...	es	?
	11 @JonaramirezX	RT @antico2: @bbimbi Desinforma bastante y promueve el dióxido de cloro ...	2020-10-11 ...	es	?
	12 @MMakedonsk...	@aacc_es Se puede curar cáncer con dióxido de cloro, peróxido de hidrógeno, vitaminas naturales(no sintéticos) y alimentos...	2020-10-11 ...	es	?
	13 @EnricBermeda...	@carmenvivomidiv @crisis_historia @A_Corrales_1020 @gemmspp @Random_Dragon @SalvaGarc4 @iPodMusica ...	2020-10-11 ...	es	?
	14 @cescept	RT @antico2: Este es el documento publicado por USA Today donde dan los argumentos por los que pidieron la detención d...	2020-10-11 ...	es	?
	15 @AlvisV	Mi papá chingándose dióxido de cloro cada día y le demostré (con fuentes confiables) que te servía para pura verga y que sól...	2020-10-11 ...	es	?
	16 @ant72_jos	@filorentantoni No es la única porque también Santiago Cúneo y Lilia Lemoine hacen apología al dióxido de cloro, o sea ha...	2020-10-11 ...	es	?
	17 @Aldeadeletras	RT @hambroides: Entrevista en Imagen Querétaro sobre el Dióxido de cloro. Parte 1...	2020-10-11 ...	es	?
	18 @caro_wylli	PANDEMIA...	2020-10-11 ...	es	?
	19 @FeFrias	RT @antico2: Este es el documento publicado por USA Today donde dan los argumentos por los que pidieron la detención d...	2020-10-11 ...	es	?
	20 @beta_foly	@cesaralva58: Algo raro. Con Ivermectina, hidroxicloroquina, dióxido de cloro, y mas yerbas, yo he visto vivir gente. Que ...	2020-10-11 ...	es	?
	21 @antico2	@FernandoCervera escuche el podcast donde hablaban de dióxido de cloro, buen trabajo	2020-10-11 ...	es	?
	22 @almadelmar33	@vivicanosak Vivi querida! Te informo, no se si sabias que la diputada Frade, pidió en una sesion de la ...	2020-10-11 ...	es	?
	23 @BebaAreta	@ttn_podemos: Dr. Alfredo Valencia: "Tomando el dióxido de cloro noté mejoría muy muy pronto, a los 6 o 7 días ya ...	2020-10-11 ...	es	?
	24 @joselan_mx	RT @lamber_2009: LA DIPUTADA ARGENTINA MONICA FRADE DEFIENDE EL DIOXIDO DE CLORO https://t.co/Ve9j3F4E7 a ...	2020-10-11 ...	es	?
	25 @joselan_mx	RT @JStrane: El dióxido de cloro está revolucionando a la medicina https://t.co/XG1q7vwVwV a través de @metrolibreptly	2020-10-11 ...	es	?
	26 @RagnaiDanne...	RT @JuanGnasmus: Encuesta...	2020-10-11 ...	es	?

Figura 2. Muestra de las columnas obtenidas por la consulta.

Fuente: elaboración propia en Orange Canvas.

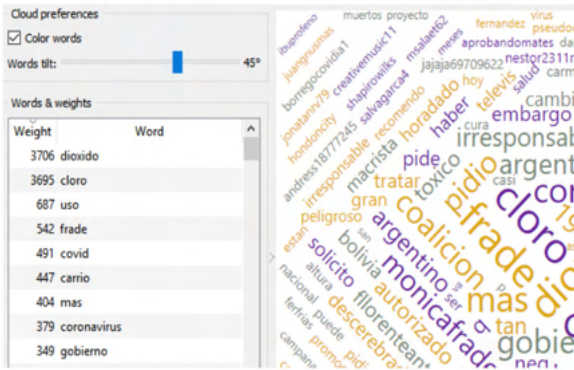


Figura 4. Word Cloud después de la limpieza de palabras.

Fuente: elaboración propia en Orange Canvas.

Después, usando el Word Cloud a simple vista se puede observar palabras más grandes que las otras, lo que indica que estas están con más frecuencia, así mismo se observa en la parte derecha Figura 4 la cantidad de veces que estas palabras se repiten como es el caso Cloro 3695 veces, Dióxido 3706 veces, uso 687 veces, Covid 497 veces, etc.

d) Modelado.

El modelo de minería de datos sobre el análisis de sentimiento basado en el lexicon sobre el uso del dióxido de cloro para el tratamiento de COVID-19 se presenta en la Figura 5.

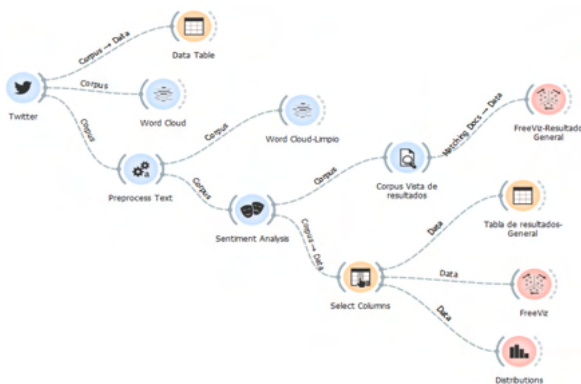


Figura 5. Diseño del modelo de minería de datos sobre el Análisis de Sentimiento en Twitter basado en lexicones sobre el uso de dióxido de cloro para el tratamiento del COVID-19.

Fuente: elaboración propia en Orange Canvas.

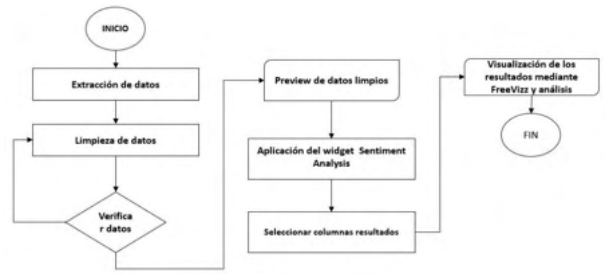


Figura 6. modelo de minería de datos sobre el Análisis de Sentimiento en Twitter basado en lexicones sobre el uso de dióxido de cloro para el tratamiento del COVID-19.

Fuente: elaboración propia.

La Evaluación y Despliegue de la metodología CRISP-DM se realiza en la parte de resultados, según el diseño del modelo que se presenta en la Figura 5.

RESULTADOS

El widget Análisis de sentimientos (Sentiment Analysis) predice el sentimiento para cada documento en un corpus. Utiliza los métodos de sentimiento Liu Hu y Vader de NLTK. Ambos están basados en léxico. Para esta investigación se usó el método Vader análisis de sentimientos basado en léxico y reglas, este módulo nos proporciona 4 tipos de resultados (positivo, neutro, negativo, compuesto) para cada documento o tweet.

Una puntuación compuesta representa el sentimiento total de un tweet, donde -1 es el más negativo y 1 el más positivo.

La puntuación compuesta se calcula sumando las puntuaciones de valencia de cada palabra en el léxico, se ajusta de acuerdo a las reglas y luego se normaliza para estar entre -1 (más extremo negativo) y +1 (más extremo positivo). Los valores umbrales típicos son:

- sentimiento positivo: compuesto ≥ 0.05
- sentimiento neutral: compuesto > -0.05 y compuesto < 0.05
- sentimiento negativo: compuesto ≤ -0.05

En cambio, los puntajes positivo, negativo y neutro representan las proporciones de texto que caen en cada categoría por los que todos deben sumar 1 o cerca de él con la operación flotante (EGilbert, 2014).

Para la visualización de los resultados mediante un gráfico y la clasificación de los mismos utilizamos el widget freeViz.

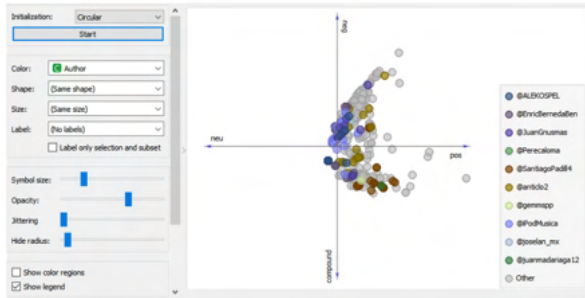


Figura 7. Vista de resultados mediante el widget FreeViz.

Fuente: elaboración propia en Orange Canvas.

Como se podrá observar en la Figura 7 se muestra el resultado de todo el análisis de sentimiento en función al autor de cada tweet, de esta manera se observa de forma general dichos resultados y por lo cual para una mejor interpretación y análisis de esta información se procederá a analizar estos resultados en función a los valores (intensidad del Tweet) obtenidos de cada tweet como negativos, positivos, neutros y compuestos.

Para ello se procedió a utilizar el widget Data sampler con el cual para la interpretación de resultados se utilizó el total de datos equivalente a 4500 Tweets.

a) Análisis de tweets con puntuación hacia positivos

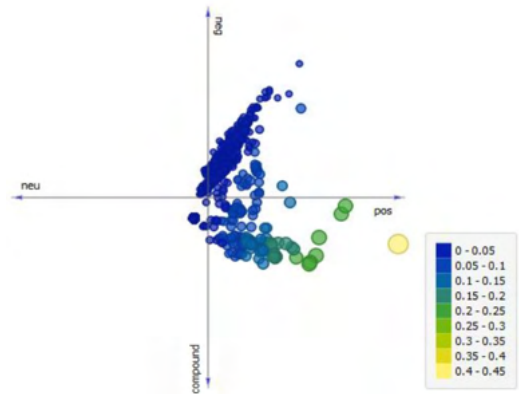


Figura 8. Resultados con puntuación de positividad con FreeViz

Fuente: elaboración propia en Orange Canvas

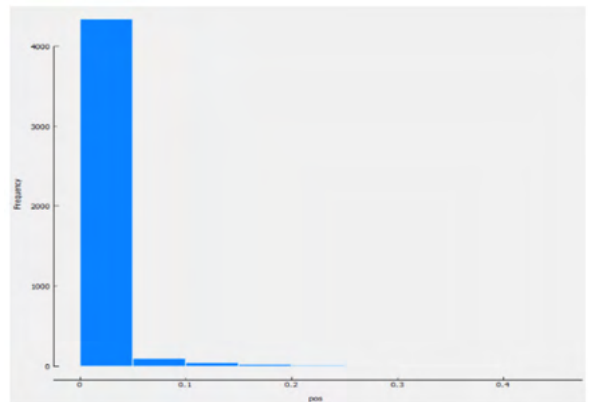


Figura 9. Gráfico de barras de resultados con puntuación hacia positivo

Fuente: elaboración propia en Orange Canvas

Como se puede observar en la Figura 9 se obtuvo en la primera barra un puntaje entre 0 y 0.05 para 4341 tweets representando al 96.47%, de igual manera la segunda barra con un valor entre 0.05 y 0.1 para 91 tweets representando al 2.02%, en la tercera barra con un valor entre 0.1 y 0.15 para 40 tweets representando al 0.89%, en la cuarta barra con un valor entre 0.15 y 0.2 para 19 tweets representando al 0.42%, en la última barra con un valor entre 0.2 y 0.25 para 7 tweets representando al 0.16%. del mismo modo en la Figura 8 se puede observar los resultados

clasificados de los tweets antes mencionados con mayor intensidad de positividad con colores más claros como el amarillo y verde, para el caso el resultado más intenso a positivo tiene un valor de 0.403. De este primer análisis podemos mencionar que, de los tweets analizados se obtuvo una intensidad positiva baja considerando que 4341 tweets de los 4500 tienen una puntuación entre los valores 0 y 0.05 que representa el 96.47% de tweets.

b) Análisis de tweets con puntuación hacia neutros

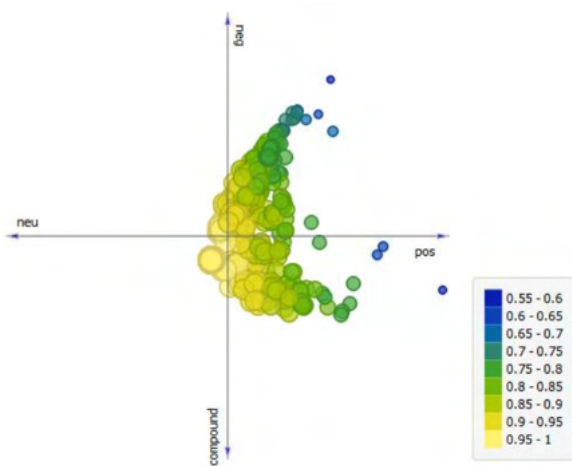


Figura 10. Resultados con puntuación de neutralidad con FreeViz.

Fuente: elaboración propia en Orange Canvas.

En la Figura 11 se muestra el análisis de la intensidad a neutralidad de los 4500 tweets; se observó que, 15 tweets se tiene valor entre 0.7 y 0.75 representando al 0.33 %, así como también 397 tweets con un valor entre 0.75 y 0.8 representando 8.82%, y 252 tweets han sido clasificados entre 0.8 y 0.85 con un porcentaje del 5.60%, 480 tweets entre 0.85 y 0.9 con un porcentaje de 10.67%, y en la penúltima barra con 446 tweets con un porcentaje de 9.91% con una intensidad de 0.9 y 0.95 y el ultima barra con 2902 tweets con un valor mayor igual a 0.95 con un porcentaje de 64.49%, de la misma manera en la Figura 10 se puede observar los resultados clasificados de los tweets antes mencionados como neutros con colores más claros y de color amarillo y verde, para nuestro análisis el resultado neutro con mayor intensidad es de 0.95. En este segundo análisis se puede mencionar que, se obtuvo 2902 tweets de los 4500 indicando una intensidad de sentimiento neutro alto, considerando que 64.49% se encuentran entre los valores de mayor o igual a 0.95.

c) Análisis de tweets con puntuación hacia negatividad

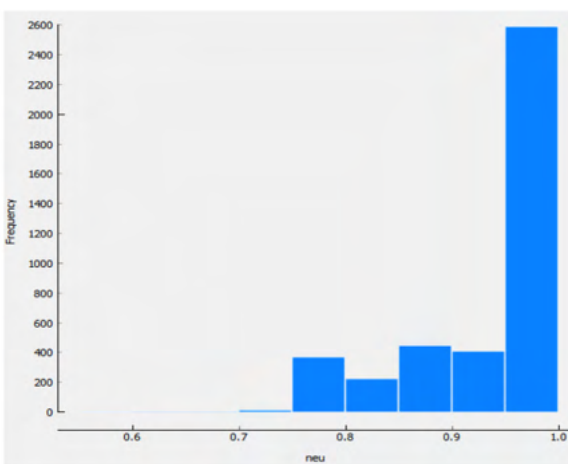


Figura 11. Gráfico de barras de resultados con puntuación hacia neutros.

Fuente: elaboración propia en Orange Canvas.

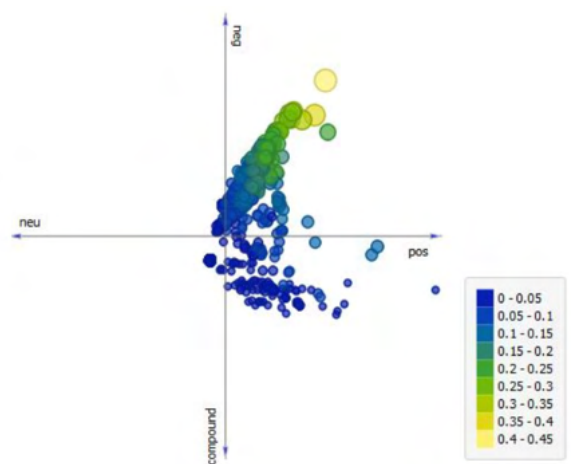


Figura 12. Resultados con puntuación de negatividad con FreeViz.

Fuente: elaboración propia en Orange Canvas.

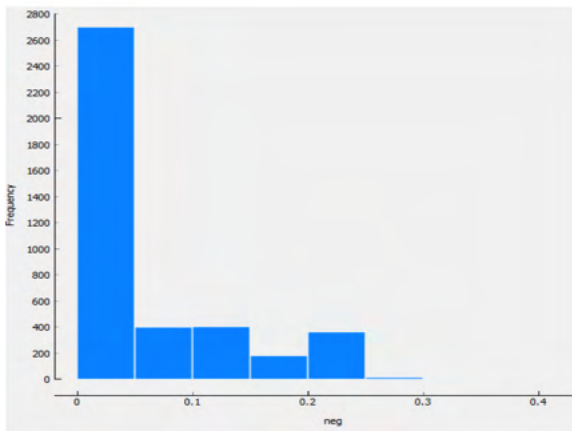


Figura 13. Gráfico de barras de resultados con puntuación hacia negatividad.

Fuente: elaboración propia en Orange Canvas.

De acuerdo a los resultados obtenidos, en la Figura 13 se muestra el análisis de la intensidad a negatividad de los 4500 tweets; la primera barra son 3018 tweets con un valor menor a 0.05 con un porcentaje de 67.07%, en la segunda barra con 435 tweets con un valor de 0.05 y 0.1 representando 9.67%, tercera barra con 436 tweets con valor entre 0.1 y 0.15 representando el 9.69%, y 205 Tweets con un valor entre 0.15 y 0.2 con un porcentaje 4.56%, y 389 tweets con un valor entre 0.2 y 0.25 con un porcentaje de 8.64%, y por último 13 Tweets con una intensidad entre 0.25 y 0.3 con un porcentaje de 0.29%, de acuerdo a estos valores podemos corroborar en la figura 12 donde se muestra el tweet con una intensidad mayor negativa de 0.423. Se tiene 2700 tweets de los 4500 cuyos valores de puntuación son menores a 0.05 representando el 67.07% indicándonos que existe una intensidad de sentimiento a -1.

d) Análisis de tweets con puntuación compuesta

La puntuación compuesta muestra la puntuación total de un comentario, cuanto más cercano a 1 sea el nivel puntuación es más positivo, de lo contrario se acerca a -1 tiende a ser más negativo. En la figura 14, se

muestra las puntuaciones compuestas, en el cual se ajustó el gráfico de barras a los valores umbrales típicos de esta manera determinar la puntuación total de los comentarios de manera unidimensional basándonos en los valores umbrales.

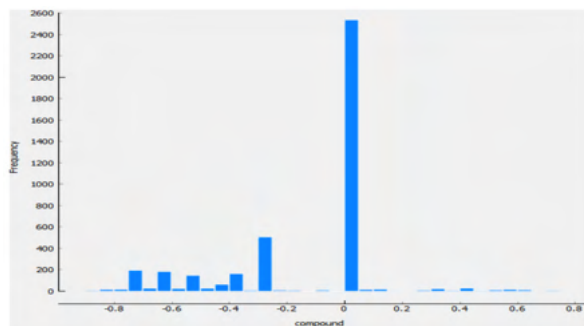


Figura 14. Gráfico de barras de valores umbrales.

Fuente: elaboración propia en Orange Canvas.

En el primer rango de la figura 14 muestra un umbral en una puntuación entre -1 y -0.05 con 1524 (33.9%) de tweets representando este rango a un sentimiento negativo en los tweets.

En los siguientes umbrales están en el rango de -0.05 y 0.05 con 2837 (63.04%) de tweets representado este rango un sentimiento neutro en los tweets. Como tercer umbral tenemos los valores entre 0.05 y 1 con 139 (3.09 %) de tweets, representando este rango como un sentimiento positivo en los tweets.

CONCLUSIONES

En la red social Twitter, de los 4500 tweets que fueron analizados, en los comentarios realizados por los usuarios acerca del dióxido de cloro como tratamiento del COVID-19, está asociado con los siguientes términos: dióxido (3706 veces), cloro (3695 veces), uso (687 veces), covid (491 veces) y coronavirus (379 veces). Los datos se obtuvieron con la palabra “dióxido de cloro” y esta a su vez vinculó a las palabras como uso, covid y coronavirus.

En los tweets de los usuarios, tomados los días 8, 9, 10 y 11 de octubre de este año 2020,

muestran sentimiento neutral y negativo sobre el dióxido de cloro como tratamiento para el COVID-19 que comentaron en la red social Twitter y de acuerdo a la investigación 1524 tweets obtuvieron una puntuación de negatividad en relación al dióxido de cloro para tratar el COVID-19 con un porcentaje de 33.9%, y 2837 tweets obtuvieron puntuación de neutralidad ante el dióxido de cloro con un porcentaje de 663.04%, y 139 tweets con una puntuación de positividad sobre el dióxido de cloro con un porcentaje de 3.09%.

REFERENCIAS

Agencia para Sustancias Tóxicas y el Registro de Enfermedades. (2004). *Dióxido de Cloro y Clorito*. Atlanta GA 30333: Agencia para Sustancias Tóxicas y el Registro de Enfermedades.

AS Perú. (26 de 07 de 2020). *Qué es y para qué sirve el dióxido de cloro y qué dice el Minsa sobre su uso*. Recuperado el 02 de 08 de 2020, de https://peru.as.com/peru/2020/07/26/actualidad/1595791000_910783.html

Cabezas Sanchez, C., Solari Zerpa, L., & Caballero Ñopo, P. (2020). Eficacia y seguridad del dióxido de cloro para el tratamiento de COVID-19. *INSTITUTO NACIONAL DE SALUD*, 2-10.

EGilbert, H. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Conference on Weblogs and Social Media (ICWSM-14)*.

Figueroa, M. G. (2008). *Búsqueda, Descubrimiento y Análisis de Sentimientos en Twitter*. Chile-Región de Valparaíso.

Mitra, S., & Acharya, T. (2005). *Data mining: multimedia, soft computing and bioinformatics*. WILWY.

Organizacion Mundial de Salud. (2020). *Preguntas y respuestas sobre la enfermedad por coronavirus (COVID-19)*. Recuperado el 05 de 07 de 2020, de https://www.who.int/es/emergencias/diseases/novel-coronavirus-2019/advice-for-public/q-a-coronavirus-es?gclid=EAIAIQobChMli-T4iYK46wIVeP7jBx1czw11EAAYASABEgKmh_D_BwE

Orgaz, C. J. (15 de 04 de 2020). *Coronavirus | Dióxido de cloro, el peligroso químico que se promociona como cura para el covid-19 y sobre el que advierten los expertos*. Recuperado el 2 de 08 de 2020, de <https://www.bbc.com/mundo/noticias-52303363>

Orihuela, J. L. (2011). *Mundo Twitter*. Navarra, España: ALIENTA.

Rubín, R. (22 de 05 de 2020). *Qué es Twitter, cómo funciona y qué te puede aportar esta red social*. Recuperado el 10 de 07 de 2020, de <https://www.ciudadano2cero.com/twitter-que-es-como-funciona/>

Trucco, F. (06 de 08 de 2020). *Bolivia aprueba dióxido de cloro como tratamiento para coronavirus*. Recuperado el 10 de 08 de 2020, de <https://cnnespanol.cnn.com/2020/08/06/camara-de-diputados-de-bolivia-aprueba-dioxido-de-cloro-como-tratamiento-para-coronavirus/>

Vallalta Rueda, J. F. (s.f.). *CRISP-DM: una metodología para minería de datos en salud*. Recuperado el 05 de 07 de 2020, de <https://healthdataminer.com/data-mining/crisp-dm-una-metodologia-para-mineria-de-datos-en-salud/>