

# Un recorrido electoral en Twitter Venezuela durante el periodo enero de 2019 a diciembre de 2020

## A electoral tour in Twitter Venezuela during the period January 2019 to December 2020

COLINA, Mairene<sup>1</sup>  
COLINA, Katherine<sup>2</sup>  
FERNANDEZ, Alberto<sup>3</sup>  
MAIA, Manuel<sup>4</sup>  
TORREALBA, Daniela<sup>5</sup>

### Resumen

Se presenta un análisis exhaustivo de cómo se percibió el tema electoral en Venezuela por parte de los usuarios de la red social Twitter. Este análisis se realizó haciendo uso de técnicas estadísticas y de inteligencia artificial que facilitan el análisis de la información disponible en grandes volúmenes de datos. Se determinaron: las opiniones de mayor impacto, el sentimiento expresado, los usuarios más activos, los más citados, etc. El lapso analizado está comprendido entre el 08-01-2019 y el 06-12-2020.

**Palabras clave:** minería texto, minería datos, twitter.

### Abstract

An exhaustive analysis of how Twitter users perceive the electoral process in Venezuela is presented. This analysis was done by using statistical techniques and artificial intelligence methods in order to simplify the analysis of large volumen of data and document collections. The most impactful opinions, the sentiment expressed, the most active and cited users, etc. were determined. The lapse analyzed is between 01-08-2019 and 12-06-2020.

**Key words:** text mining, data mining, twitter.

---

## 1. Introducción

Como parte del trabajo de observación electoral, la Organización No Gubernamental Red de Observación Electoral de la Asamblea de Educación (ROAE, [www.redobservacion.org](http://www.redobservacion.org)) debe estar atenta a la opinión de la población en torno a los eventos electorales, políticos y sociales que ocurren en Venezuela, en particular en un ambiente altamente polarizado como el que se vive en los últimos años. Es por ello que la ROAE, a través de su equipo técnico, se planteó realizar una revisión de los contenidos de los tweets teniendo como punto de atención

---

<sup>1</sup> Profesora. Escuela de Matemática. Universidad Central de Venezuela. Venezuela. Email: mcolina@gmail.com

<sup>2</sup> Estudiante. Escuela de Computación. Universidad Central de Venezuela. Venezuela. Email: kathymcolina@gmail.com

<sup>3</sup> Profesor. Escuela de Química. Universidad Central de Venezuela. Venezuela. Email: ferna.aj@gmail.com

<sup>4</sup> Profesor. Escuela de Matemática. Universidad Central de Venezuela. Venezuela. Email: manuelpmaiaf@gmail.com

<sup>5</sup> Profesora. Escuela de Matemática. Universidad Central de Venezuela. Venezuela. Email: danielatorrealbag@gmail.com

los procesos electorales en marcha y la reacción del ciudadano en respuesta a las situaciones, decisiones u otros que se generen en torno a ese proceso electoral.

Las redes sociales, en particular Twitter, se han convertido en una vía de expresión y opinión. En Venezuela es quizás la herramienta más importante para enterarse de situaciones que suceden a diario y expresar opiniones con cierta libertad. Los diversos estudios que se han generado en torno al valor de esta red como fuente de información y los diversos trabajos de investigación que sobre los datos emanados por ella se han hecho, por ejemplo, para el estudio de procesos migratorios como hacen en Gualda y Rebollo (2020) o para análisis de la comunicación en situaciones de emergencia como en Suau-Gomila, Percastre, Palà y Pont (2017), dan cuenta del valor de esta red como fuente de datos.

Para realizar este estudio se descargaron diariamente, mediante el uso de una interfaz de programación de aplicaciones (API) de twitter y códigos desarrollados en los lenguajes de programación Python y R, los tweets que contenían el tema electoral venezolano en el lapso comprendido entre el 08-01-2019 y el 06-12-2020 y, posteriormente, se analizó su contenido, haciendo uso de técnicas estadísticas y de inteligencia artificial que facilitan el análisis de la información disponible en grandes volúmenes de datos y colecciones de documentos. Dichas técnicas son conocidas como minería de datos y minería de texto. Para mayor información sobre estas técnicas, ver Jo (2019).

Con la ayuda de las herramientas disponibles de la minería de datos y de texto se pueden Identificar “hechos” y datos puntuales a partir de los tweets, agrupar tweets similares, determinar los temas tratados en los tweets mediante la categorización automática de los textos, visualizar las relaciones entre los conceptos tratados en los tweets, etc. Además, se puede realizar análisis de sentimientos lo que también se conoce como minería de opinión.

Para analizar el conjunto total de datos recopilados (un total de 126375 tweets), el lapso total de estudio (2019-2020) se dividió en seis periodos de cuatro meses cada uno:

- Periodo 1: enero-abril 2019.
- Periodo 2: mayo-agosto 2019.
- Periodo 3: septiembre-diciembre 2019.
- Periodo 4: enero-abril 2020 .
- Periodo 5: mayo-agosto 2020.
- Periodo 6: septiembre-diciembre 2020.

Una vez pre-procesados los tweets, se realizó un análisis exhaustivo en el que se identificaron los días en los cuales el tema electoral motivó a un número mayor de usuarios a emitir o difundir opiniones respecto a este tema. También se identificaron:

- a) los tuiteros con más seguidores y los mensajes que éstos difundieron en esos días de más impacto,
- b) los hashtags y los tweets que recibieron más difusión y
- c) los diferentes temas abordados por cada periodo analizado.

Finalmente se hizo un análisis sobre cómo han percibido los usuarios de esta red el tema electoral, en lo que denominamos análisis de sentimientos.

---

## 2. Metodología

En este capítulo se describe la metodología usada. Esto incluye las herramientas disponibles en la red, como las API de Twitter, para, en primer lugar, extraer la información de interés. También se usaron ampliamente algunas herramientas e ideas propias de la minería de texto, como los bigramas y el análisis de sentimientos. Para todo el manejo y procesamiento de la información se hizo uso de los lenguajes de programación Python y R, utilizando las facilidades de sus librerías y desarrollando códigos propios para fines específicos.

### 2.1. Extracción de la información

Twitter tiene a disposición de los desarrolladores una API de búsqueda que permite obtener tweets publicados desde marzo de 2006 hasta la actualidad (<https://help.twitter.com/es/rules-and-policies/twitter-api>). Una API (interfaz de programación de aplicaciones, por sus siglas en inglés) es un mecanismo para conectar dos softwares y realizar intercambio de datos entre ellos, en este caso entre Twitter y el software desarrollado para esta investigación. Entre las funcionalidades de la modalidad Premium de esta interfaz, están la descarga de datos en vivo y de cuentas específicas. La recopilación de los tweets de interés para esta investigación comenzó con la creación de una cuenta de desarrollador en Twitter para poder acceder a la API, y de esta manera, con las claves suministradas por la API, se extrajeron los tweets que conforman la data analizada.

La API Premium retorna un máximo 100 tweets por solicitud, con un límite de 180 solicitudes cada 15 minutos. Además, para una cuenta de usuario particular, el límite son los 3200 tweets más recientes. Hay que destacar que la búsqueda Premium requiere de autorización por parte de la API.

Para las búsquedas existen reglas que se componen de una o más cláusulas, una palabra clave o una frase exacta. Se puede combinar varias cláusulas con operadores lógicos “y” y “o”.

En el caso del operador Premium (usado en este proyecto), cada regla puede tener hasta 2048 caracteres de longitud sin límites en el número de cláusulas positivas. Además, se pueden usar coordenadas de geolocalización para buscar mensajes emitidos desde una región determinada. En este caso se utilizaron las coordenadas de Venezuela.

Para la extracción de la información y la búsqueda específica de palabras claves se usó la librería `searchtweets` de Python conectando ésta a la API de Twitter. Previamente se diseñó un código en Python que permite darle formato de tabla a la información extraída para posteriormente analizar esta información en R.

Para la descarga de tweets de cuentas específicas se diseñó un código en R. Para poder acceder a los tweets de las cuentas de interés se hizo uso de la librería `rtweet` conectándola a la API de Twitter.

### 2.2. Pre-procesamiento de la información

Un paso fundamental antes de iniciar el análisis de cualquier texto es el pre-procesamiento. En esta etapa se procede a realizar una limpieza de los datos y uniformizar la información. En el contexto de las redes sociales esto resulta fundamental pues abundan las faltas de ortografía.

El pre-procesamiento incluye la eliminación de los acentos, signos de puntuación, caracteres especiales tales como el asterisco, llamados a páginas web, etc, así como la extracción de las palabras separadas entre sí por signos de puntuación. Se eliminan también las preposiciones, artículos, conjunciones, etc, pues estas no aportan información. Cada tweet que se va a procesar se convierte en formato de texto plano. El conjunto de tweets conforman el corpus el cual es posteriormente analizado.

## 2.3. Procesamiento de la información (minería de texto y datos)

En este subcapítulo se describen los procesos y métodos utilizados para analizar la información una vez pre-procesada.

### 2.3.1 Cronología de eventos noticiosos

El análisis de la cronología de los eventos noticiosos resaltantes relativos al tema electoral es un componente importante, pues la visualización de ésta combinada con el análisis de frecuencia se puede utilizar para identificar qué fenómeno o situación causó mayor actividad en la red social.

La data extraída cuenta con la fecha de emisión de cada uno de los mensajes. En primer lugar, la fecha se segmenta en día, mes y año. Posteriormente, se contabiliza la cantidad de mensajes por periodo, mes y día. Esto permite identificar, para cada periodo en estudio, tanto el mes como el día de mayor actividad. Finalmente se procede a determinar, las noticias más resaltantes y, que por tanto, pudieron ser las de mayor impacto y difusión en la red social Twitter esos días de mayor actividad.

### 2.3.2. Identificación de *hashtags* (etiquetas) y *arrobos* (menciones)

Una vez identificado el día con mayor actividad, es interesante conocer qué información trataron de divulgar los usuarios con más seguidores y la cantidad de veces que dicho mensaje fue replicado. Esta metodología permite identificar en términos generales cual fue el tópico principal difundido.

El uso del símbolo *hashtag* (#) antes de una palabra o frase permite categorizar los tweets y ayudarlos a mostrarse más fácilmente en la búsqueda de Twitter. Los usuarios suelen hacer el uso del *hashtag* de dos formas, para tratar de colocar en tendencia un mensaje o información particular o usan un *hashtag* que ya es tendencia para tratar de divulgar su mensaje. En el análisis de Twitter buscamos el símbolo # y con él extraemos la frase o palabra que lo sigue. Una vez identificados los *hashtags*, se contabiliza el total de veces que cada uno fue implementado, para finalmente identificar aquellos que lograron mayor difusión.

Otro punto interesante en el análisis es conocer qué usuario fue el más mencionado. Esto se hace identificando el símbolo @ en el texto y posteriormente se extrae el texto que lo acompaña.

### 2.3.3. Bigramas

Al tratar las palabras de un texto como unidades independientes, se pueden identificar las palabras más frecuentes y de esta manera realizar algunos análisis sencillos. Sin embargo, esto no permite dar cuenta del orden de las palabras ni de las frases presentes en los textos analizados. En este sentido, los  $n$ -gramas son muy útiles pues permiten dividir el texto en segmentos de  $n$  palabras continuas (también llamados tokens), logrando de esta manera capturar muy de cerca el contexto. Cada uno de estos cortes o divisiones del texto se llama  $n$ -grama. Así, el corpus se convierte en una colección de  $n$ -gramas y cada documento en el corpus es representado por un vector. Este vector contiene la frecuencia de cada  $n$ -grama presente y cero para los  $n$ -gramas que no están presentes. Al ver la frecuencia con la que una palabra dada ( $X$ ) va seguida de una palabra dada ( $Y$ ), podemos construir un modelo de las relaciones entre ellas. Este modelo estadístico además asigna probabilidades a las secuencias de las palabras.

En este trabajo se considera  $n = 2$ , lo cual genera bigramas que permiten resumir y analizar la información en los textos extraídos. Este modelo además de proporcionar la frecuencia de los pares de palabras en cada tweet, aproxima la probabilidad condicional de una palabra, dadas todas las palabras anteriores,  $P(w_n | w_1, w_2, \dots, w_{n-1})$ , usando solamente la probabilidad condicional de la palabra dada la palabra anterior  $P(w_n | w_{n-1})$ . Para el análisis e interpretación de la información, una vez identificados los pares de palabras más

frecuentes, los bigramas son graficados aplicando técnicas de grafos y las librerías apropiadas de R (ver Silge y Robinson (2017)).

#### **2.3.4. Análisis de sentimientos**

El análisis de sentimientos consiste en analizar las emociones que subyacen en un texto determinado. Esto ayuda a comprender la opinión y los sentimientos en un texto y clasificarlos como positivos, negativos o neutrales. Esta técnica permite entre otras cosas analizar publicaciones en redes sociales, lo cual se conoce como minería de opinión.

Hay dos tipos principales de opiniones: opiniones regulares y opiniones comparativas. Una opinión comparativa expresa una relación de similitudes o diferencias entre dos o más entidades. Por su parte una opinión regular es simplemente un sentimiento, una actitud, una emoción o una valoración positiva o negativa sobre una entidad o de un aspecto de la entidad por parte de un titular de opinión. Lo positivo, negativo y neutral se denominan orientaciones de opinión, también llamadas orientaciones de sentimiento u orientaciones semánticas.

Las palabras de opinión positiva se usan para expresar algunos estados deseados, mientras que las palabras de opinión negativas se usan para expresar algunos estados no deseados. Ejemplos de palabras de opinión positivas son: bueno, hermoso, agradable. Ejemplos de palabras de opinión negativa son: malo, porquería y terribles. Además de las palabras individuales, también hay frases de opinión y expresiones idiomáticas, por ejemplo, “le cuestan a alguien un brazo y una pierna”. Colectivamente, se les llama el léxico de opinión. Son fundamentales para la minería de opiniones por razones obvias. Existen diversas técnicas para analizar los textos de opinión. Una de las más implementadas está basada en un diccionario de palabras con ponderaciones. Para más detalles sobre la minería de opinión ver Pang y Lee (2008).

En este trabajo se implementó un diccionario usado internacionalmente contenido de una lista de palabras y frases las cuales tienen asignado un valor o ponderación en el rango de valores de -5 a 5. La ponderación asignada demarca la intencionalidad del sentimiento. Los valores menores que cero son asignados a las palabras de opinión negativa y los valores mayores de cero son asignados a las palabras de opinión positiva. Dicho diccionario fue ampliado con palabras usadas en el argot venezolano. Como parte del procedimiento, cada texto, o tweet, es segmentado en palabras y éstas, con la ayuda del diccionario, son debidamente categorizadas. La palabra "no" es eliminada de la lista de palabras ya que esta es una palabra muy común en español que no necesariamente implica un sentimiento negativo.

Posteriormente, para cada texto se realiza la suma de los pesos o ponderaciones asociadas a cada palabra, con lo cual cada tweet obtiene un puntaje. Si el puntaje está por encima de cero, el tweet es clasificado como positivo y si la puntuación es por debajo de cero, el tweet es considerado negativo. En caso de que la puntuación sea cero el tweet es considerado neutro. Finalmente, para tener una visión de cómo los venezolanos percibieron día a día el tema electoral, se procedió a promediar, por día, la puntuación de los tweets de la muestra. De esta manera tenemos una visión promedio por día del sentimiento expresado en los tweets.

Otra metodología empleada, además de la ya descrita, con la que se logra extraer tendencias a nivel de sentimientos, es el uso de regresiones locales aplicando el método Loess. Este método permite generar una curva que aproxima el comportamiento del sentimiento general presente en las opiniones en estudio a lo largo de todo el periodo considerado. Para aplicar este método, se generó un entorno, en función de los puntos vecinos más cercanos, de cada punto de la variable dependiente, que en este caso es el valor promedio del sentimiento diario, y se ajustó a un polinomio utilizando mínimos cuadrados ponderados, dando más peso a los puntos cercanos al punto cuya respuesta se está estimando y menos peso a los puntos más alejados.

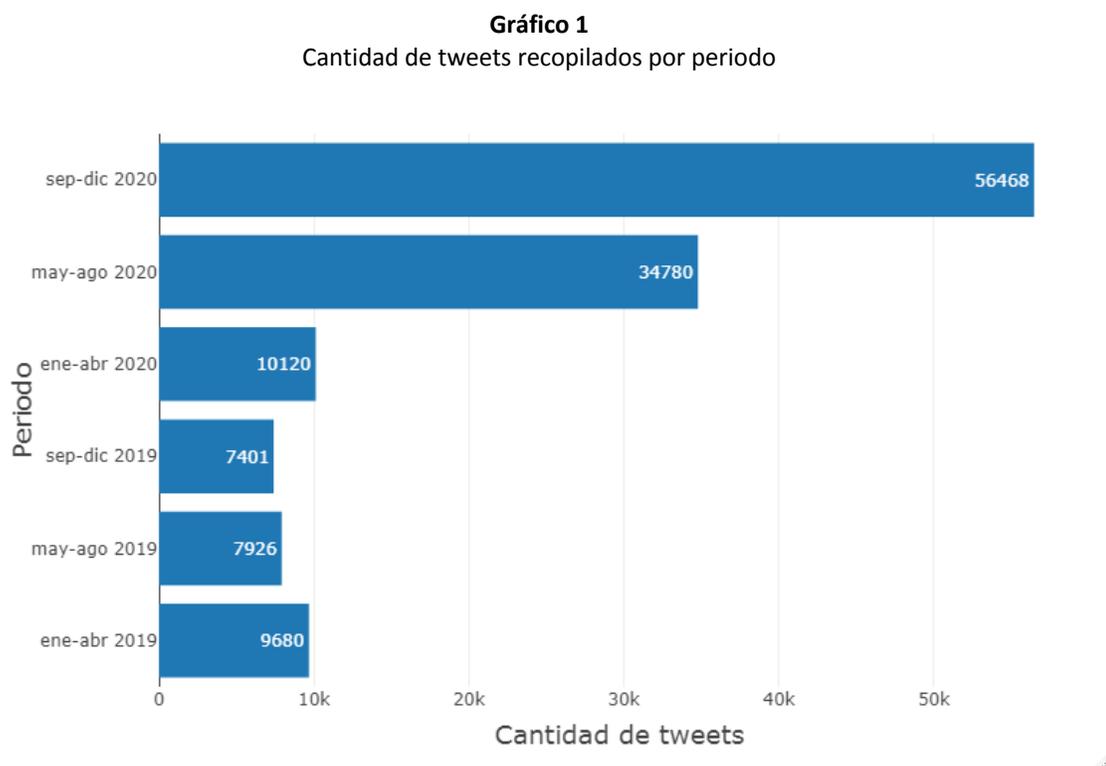
### 3. Resultados y discusión

En este capítulo se presentan los resultados obtenidos al analizar el conjunto de tweets recopilados. En primer lugar el estudio determinó los meses y días en los cuales los usuarios se mostraron más activos emitiendo opiniones respecto al tema electoral, identificó los usuarios con mayor cantidad de seguidores y también determinó los mensajes de estos usuarios que alcanzaron mayor difusión en los días de mayor actividad en la red. Seguidamente, mediante el uso de bigramas, se logró determinar los principales temas de conversación presentes en los textos de los tweets. Finalmente, se hace un análisis mucho más detallado del conjunto de los 126375 tweets analizados, dando una mirada más detallada a los usuarios que mostraron más actividad, las tendencias presentes en los tweets y se hizo un análisis de los sentimientos promedio expresados en los textos de los tweets.

#### 3.1. ¿Qué motiva el tema electoral en la red social Twitter?

En la búsqueda de indicios e información precisa que ayuden a entender qué motiva la discusión del tema electoral en la red social Twitter, basados en la muestra total recolectada durante estos casi dos años de información, en una primera aproximación se enfocó en determinar los meses y los días en los cuales los usuarios de esta plataforma fueron más activos al emitir sus opiniones respecto al tema considerado. Además, se identificaron los usuarios con mayor cantidad de seguidores (dentro de la muestra), pues en principio sus mensajes llegan a más personas, y finalmente se determinaron los mensajes de estos usuarios que alcanzaron mayor difusión en los días más movidos en la red.

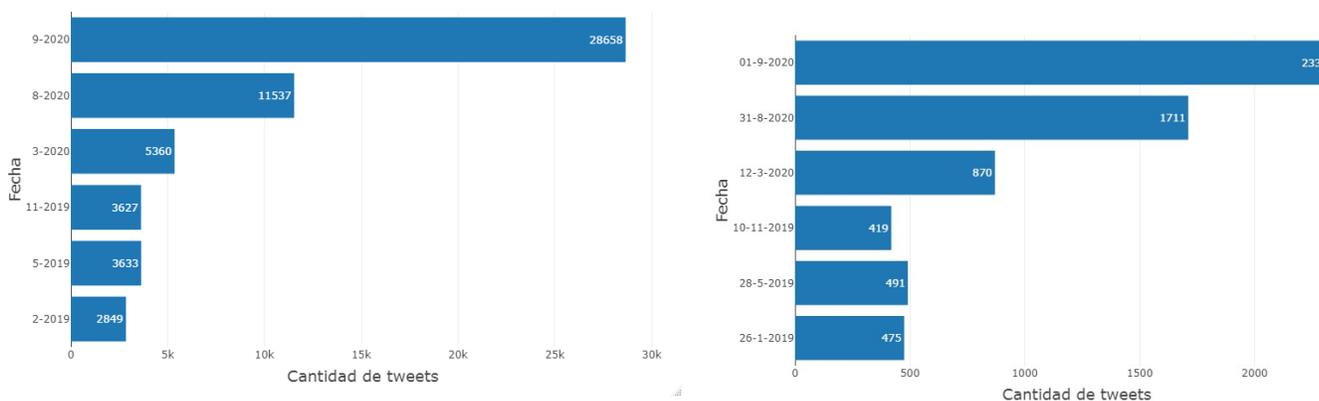
A continuación se muestran (ver gráfico 1) la cantidad de tweets recopilados en cada uno de 6 periodos. Como puede apreciarse, el periodo en el cual los usuarios opinaron en mayor cuantía respecto al tema electoral en Venezuela fue el de septiembre-diciembre de 2020.



Fuente: Los autores.

Por otra parte, en el gráfico 2, se presenta: a la izquierda los meses de mayor actividad y a la derecha los días de mayor actividad, para cada periodo. Como se puede apreciar, el mes de septiembre de 2020 fue el mes de mayor actividad, siendo el día 1/09/2020 el de mayor actividad.

**Gráfico 2**  
Meses y días de mayor actividad para cada periodo en estudio.



Fuente: Los autores.

A la izquierda del gráfico 2 se aprecia que para el primer periodo, enero-abril de 2019, el mes en el cual se habló más sobre el tema electoral fue febrero de 2019, con 2849 tweets. Sin embargo, como puede verse a la derecha del gráfico 2, el día que más tweets se recolectaron en los cuatro meses de ese periodo fue el 26 de enero de 2019, con 475 tweets recopilados. Ese día se conoció la solicitud de España a la Unión Europea de reconocer a Juan Guaidó (presidente de la Asamblea Nacional) como presidente de la República, si Nicolás Maduro no convocaba elecciones.

Un análisis detallado permitió determinar, entre los tweets recolectados el día 26 de enero de 2019, el emitido por el usuario con más seguidores y que obtuvo más retweets, así como el tweet que obtuvo más retweets ese día (ver tabla 1).

**Tabla 1**  
Tweets emitidos por los usuarios con más seguidores el 26 de enero de 2019.

Usuario	Seguidores	Retweets	Texto
@UNoticias	3919403	17	#AHORA Canciller Arreaza: hagan elecciones en el Reino Unido, metáanse en sus asuntos. <a href="https://t.co/ch9BRzBZlv">https://t.co/ch9BRzBZlv</a>
@AmalCandanguera	1116	403	A Guaidó le quedarían 27 días para hacer elecciones si estuviese en un escenario real y constitucional de falta absoluta, según el 233. Ah perdón, cierto que es una simulación. Habrán elecciones en el 2025.

Fuente: Los autores.

Para el segundo periodo, mayo-agosto de 2019, el mes de mayor actividad fue el mes de mayo de 2019, en el cual la muestra recolectada fue de 3633 mensajes, y el día en el cual los usuarios de la red emitieron más opiniones respecto al tema en estudio fue el 28 de mayo (ver gráfico 2). La noticia destacada ese día 28 de mayo, fue la reunión en Oslo (Noruega) entre representantes del gobierno y representantes de Juan Guaidó.

El tweet emitido por el usuario con más seguidores así como el que obtuvo más retweets el día de mayor actividad en el segundo periodo, 28 de mayo, se muestran en la siguiente tabla.

**Tabla 2**

Tweets emitidos por los usuarios con más seguidores el 28 de mayo de 2019.

Usuario	Cantidad de seguidores	Texto	Cantidad de retweets
@unidadvenezuela	1218624	#ATENCIÓN Circula falso documento sobre una supuesta agenda y/o acuerdo en Noruega. Es absolutamente falso. Ratificamos la ruta planteada por el Presidente de la República Guaidó: -Cese de la Usurpación -Gobierno de Transición -Elecciones libres.	0
@ReporteYa	684971	#Z28May Circula falso documento sobre una supuesta agenda y/o acuerdo en Noruega. Es absolutamente falso. Ratificamos la ruta planteada por el Presidente (E) de la República, .@jguaido: Cese de Usurpación Gobierno de Transición Elecciones libres. - @Presidencia_VE	111

Fuente: Los autores.

En el tercer periodo, septiembre-diciembre de 2019, el mes de mayor actividad corresponde a noviembre y el día que más tweets se recolectaron en estos cuatro meses fue el 10 de noviembre (ver gráfico 2). Las noticias destacadas ese día, están asociadas al ámbito internacional. A saber, la renuncia del presidente de Bolivia, Evo Morales, luego de la presentación de un informe de la OEA donde se revelaban presuntas irregularidades en los comicios realizados en octubre donde éste participó como candidato presidencial, y las elecciones generales de España. Por esta razón los venezolanos usuarios de la plataforma Twitter, hablaron de estos dos eventos, pero sin perder oportunidad de hacer comparaciones con procesos electorales o situaciones políticas en Venezuela.

La siguiente tabla (tabla 3) muestra los tweets emitidos por el usuario con más seguidores, que obtuvo más retweets, y el tweet que obtuvo mayor retweets el día 10 de noviembre de 2019.

**Tabla 3**

Tweets emitidos por los usuarios con más seguidores el 10 de noviembre de 2019.

Usuario	Cantidad de seguidores	Texto	Cantidad de retweets
@ReporteYa	736770	#10Nob - Evo Morales anuncia nuevas elecciones presidenciales -- OEA demostró que hubo FRAUDE ELECTORAL - No tiene respaldo popular, ni policial, ni militar Evo Morales debe renunciar a su cargo por delito electoral, por llenar de sangre las calles de Bolivia @luisgonzaloprz	77
@greyesg	95596	Como @evoespueblo desconoció un referéndum que le impedía reelegirse y se robó las elecciones donde nunca tuvo que haber participado, y el Pueblo salio a la calle a visibilizar su indignación, ahora quieren repetir las elecciones!! Eso no existe!! Evo, preso y Mesa, a gobernar!!	106

Fuente: los autores.

En el periodo, enero-abril de 2020, se observa un considerable incremento en el tema electoral durante el mes de marzo, obteniéndose un valor superior a los 5000 tweets para ese mes (ver gráfico 2). Se pudo observar que el día que más tweets se recolectaron en estos cuatro meses fue el 12 de marzo de 2020 (ver gráfico 2). Las noticias destacadas el 12 de marzo de 2020 fueron el incendio de los galpones (almacenes) del Consejo Nacional Electoral (CNE) y la convocatoria hecha por Juan Guaidó a una movilización ese día hacia el Palacio Federal

Legislativo con el objetivo de entregar un pliego nacional de conflicto. El tweet emitido por el usuario con más seguidores y el tweet que obtuvo más retweets recolectados ese día se muestran en la tabla 4 a continuación.

**Tabla 4**  
Tweets emitidos por los usuarios con más seguidores el 12 de marzo de 2020.

Usuario	Cantidad de seguidores	Texto	Cantidad de retweets
@ibepacheco	1913579	Según la versión del Ministro de Interior, Néstor Reverol, el incendio ocurrido en los galpones del CNE en filas de Mariche fue un acto terrorista y de sabotaje al sistema electoral. Asegura que fue intencional y que usaron combustible para acelerar la propagación del fuego.	0
@NestorLReverol	391108	Por instrucciones del Presidente @NicolasMaduro hemos presentado las pruebas del acto terrorista realizado contra el #CNE el pasado #7Mar en el Galpón ubicado en Filas de Mariche #Miranda, donde se encontraba parte de la maquinaria electoral del país #12Mar <a href="https://t.co/uwC0Um7MgJ">https://t.co/uwC0Um7MgJ</a>	453

Fuente: los autores.

Para el periodo mayo-agosto de 2020, el mes de agosto, además de ser el que presentó mayor actividad en el periodo, resultó ser aquel donde la plataforma fue usada por más usuarios que en los lapsos anteriores para dar su opinión respecto al ámbito electoral, alcanzando números superiores a los 11000 tweets, como puede observarse a la izquierda del gráfico 2. El día que más tweets se recolectaron en estos cuatro meses, como se observa a la derecha del gráfico 2, fue el 31 de agosto de 2020. La noticia destacada ese día 31 de agosto de 2020, en lo que concierne al tema de análisis, fue el anuncio del decreto presidencial donde se otorga el indulto a más de cien presos políticos, presos o imputados en el exilio.

Los tweets emitidos por el usuario con más seguidores el 31 de agosto de 2020, se muestran en la tabla 5.

**Tabla 5**  
Tweets emitidos por los usuarios con más seguidores el 31 de agosto de 2020.

Usuario	Cantidad de seguidores	Texto	Cantidad de retweets
@GenPenaloza	672449	Anuncian indulto presidencial masivo para presos políticos. Entrará en efecto luego de ser publicado en gaceta. Posiblemente este ligado a las negociaciones para elecciones. Cien personas entre diputados y periodistas serán beneficiados por la medida. <a href="https://t.co/N4k6vvcfk4">https://t.co/N4k6vvcfk4</a>	0
@Taniatabares17	5518	Soy chavista al 1000% pero hay cosas que no me parecen y lo tengo que decir camaradas, esta debe ser la dinámica, voto por la revolución, vivo en mi patria 🇨🇵 y si estoy en desacuerdo democráticamente puedo hacer llegar mi voz en desacuerdo @NicolasMaduro @dcabellor 🍷💖	223

Fuente: los autores.

Para el periodo septiembre-diciembre de 2020 se observó que el mes de mayor actividad fue septiembre y el día de mayor actividad, el primero de ese mismo mes (ver gráfico 2). Las noticias destacadas ese día, 1 de septiembre de 2020, fueron la promoción de las elecciones parlamentarias por parte de Nicolás Maduro, anunciando además cambios en los ministerios y promoviendo la participación en estas elecciones de los ministros salientes, y la nota

de prensa del CNE donde se presenta una “nueva solución tecnológica avanzada” (máquina de votación) para aplicarse en las elecciones del 6 de diciembre de 2020.

Un análisis detallado permitió determinar, entre los tweets recolectados día 1 de septiembre de 2020, los que consideramos de mayor difusión, el emitido por el usuario con más seguidores y el que obtuvo más retweets. Estos tweets pueden verse en la tabla 6.

**Tabla 6**  
Tweets emitidos por los usuarios con más seguidores el 1 de septiembre de 2020.

Usuario	Cantidad de seguidores	Texto	Cantidad de retweets
@teleSURtv	1928365	#Venezuela   #CNE implementará tecnología de avanzada en los comicios a realizarse el 6 de diciembre. #Septiembre2020 #BienvenidoSeptiembre #1Septiembre <a href="https://t.co/ZVt5kk93XK">https://t.co/ZVt5kk93XK</a>	0
@Taniatabares17	5530	Ya con este mensaje me despido, no dejo de estar molesta, pero tengamos en cuenta de que nuestro deber patrio es ir a sufragar este 6D por nuestros candidatos de la izquierda, a la derecha ni agua, es mi opinión. Chávez Vive Carajo hoy más que nunca 🥰❤️ <a href="https://t.co/G1sTMylJ5M">https://t.co/G1sTMylJ5M</a>	165

Fuente: los autores.

### 3.2. Revisión de la data. Una mirada exhaustiva sobre las opiniones respecto al tema electoral y los tópicos que lo acompañaron entre 2019 y 2020

Junto al asunto electoral, a diario surgen temas de menor impacto informativo, que son de interés para el análisis en las redes sociales y, particularmente en Twitter, ya que reflejan buena parte de las preocupaciones del venezolano en torno al desarrollo de su cotidianidad. Para conocer estos temas se aplicó una mirada más exhaustiva sobre estas opiniones.

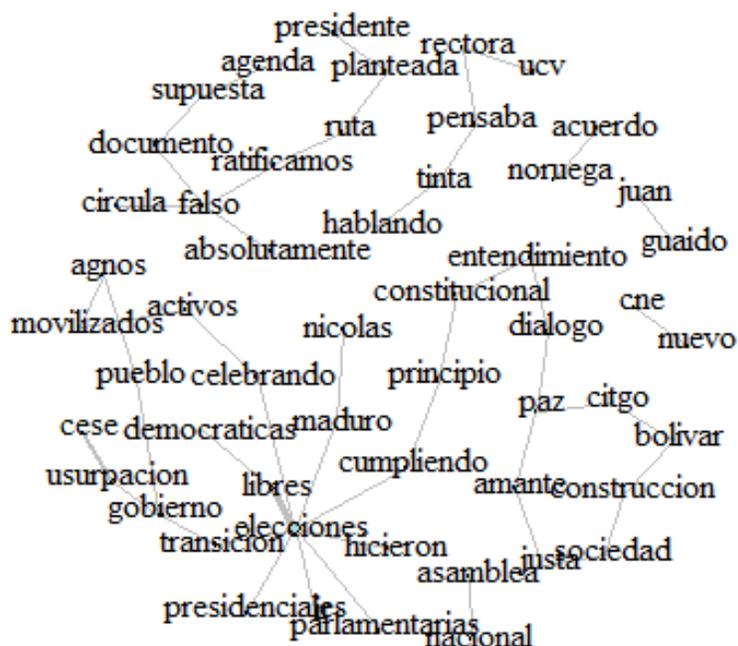
Para analizar los temas a los cuales se hace referencia en los textos de los tweets se generaron bigramas con los cuales se determinó estadísticamente los principales temas de conversación en los 126375 tweets descargados. Cada par de palabras se representó en un plano, conectadas por una flecha y, en el caso que una misma palabra se emplea acompañada de otras, se genera una especie de estrella. También se generan cadenas si una palabra se repite en otro par y así sucesivamente con otra. Este proceso se hizo para cada periodo tomando en consideración que los pares de palabras se repitieran al menos 120 veces. Los resultados obtenidos se muestran a continuación para cada uno de los periodos.

En el gráfico de bigramas para el primer cuatrimestre de 2019 (ver gráfico 3) se puede apreciar como la palabra “elecciones” es el centro de una estrella, de puntas de diferentes longitudes. También se observan cadenas aisladas de diferente longitud, con secuencias de palabras.



**Gráfico 4**

Bigrama para el periodo  
mayo-agosto de 2019



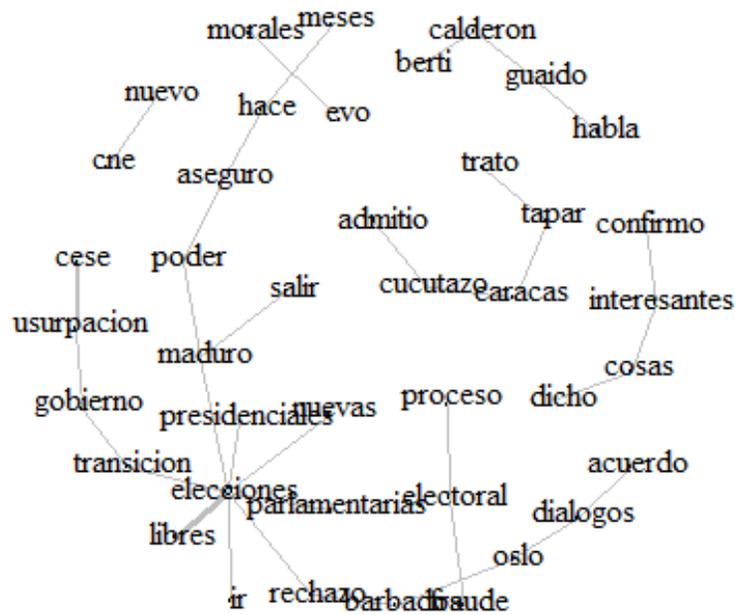
Fuente: Los autores

En el gráfico 4 se puede ver una continuidad en el tema asociado al discurso del presidente de la Asamblea Nacional Juan Guaidó: cese de la usurpación, gobierno de transición y elecciones (presidenciales) libres, así como la creación de un nuevo CNE. Otros tópicos sobre los cuales los usuarios de la plataforma opinaron fueron:

- El documento, desmentido por el centro de comunicaciones de Juan Guaidó, que circuló en las redes donde se mencionaba un acuerdo firmado en Noruega entre representantes de Juan Guaidó y de Nicolás Maduro.
- La Asamblea Nacional.
- La tinta indeleble que se usa en los procesos electorales para evitar fraudes.
- El discurso asociado al partido de gobierno: amantes de la paz, gobierno y pueblo movilizados por años para la construcción de una sociedad justa.
- El presidente Nicolás Maduro.

Analicemos ahora el bigrama del último cuatrimestre de 2019 (ver gráfico 5).

**Gráfico 5**  
Bigrama para el periodo  
septiembre-diciembre de 2019



Fuente: Los autores

Notamos (ver gráfico 5) que en los últimos cuatro meses del 2019 aún seguía vigente en la opinión de la ruta promovida por el presidente de la Asamblea Nacional Juan Guaidó: cese de la usurpación, gobierno de transición y elecciones libres. Pero ya para este periodo se insistía concretamente en elecciones presidenciales. También seguía presente el tema de la creación de un nuevo CNE. Para estas fechas además opinaron sobre:

- La destitución de Humberto Calderón Berti como representante diplomático en Colombia del presidente de la Asamblea Nacional Juan Guaidó, con motivo del llamado “Cucutazo”.
- Fraude electoral y Evo Morales. Hay que recordar que entre octubre y noviembre de 2019 se desarrollaron protestas en Bolivia luego de que se acusara al entonces presidente Morales de haber cometido fraude en las elecciones presidenciales realizadas por esos días.
- El diálogo entre el gobierno y la oposición en el marco del mecanismo de Oslo.

Para el primer cuatrimestre de 2020 el gráfico de bigramas obtenido se muestra a continuación (gráfico 6).



**Gráfico 7**

Bigrama para el periodo  
mayo-agosto de 2020



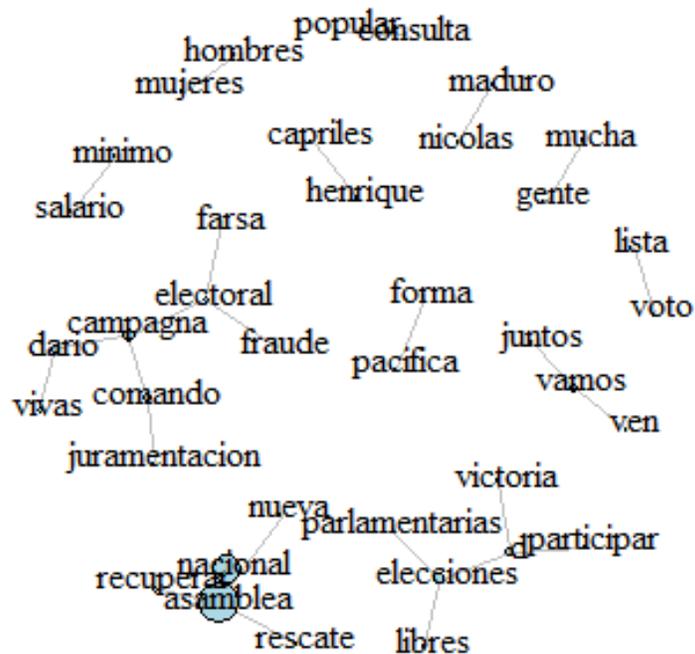
Fuente: Los autores

En el gráfico 7 se puede ver que el tema o ruta asociado al discurso del presidente de la Asamblea Nacional aún seguía en boga, junto a los siguientes temas:

- La polémica motivada por la aprobación del Reglamento Especial que regirá la elección de los representantes indígenas a la Asamblea Nacional, donde la votación directa pasa a ser un sistema de delegados que eligen en asambleas.
- El rector del CNE Rafael Simón Jiménez quien defendió el Reglamento Especial que regirá la elección de los representantes indígenas a la Asamblea Nacional.
- Elecciones Parlamentarias.
- Las elecciones, vistas como una farsa por algunos sectores de la oposición.
- Los presos políticos.
- La mención de algunos actores políticos como Claudio Fermín, Javier Bertucci, Timoteo Zambrano, Felipe Mujica, entre otros.
- Modificación de la Constitución.

Para el último cuatrimestre de 2020 crecieron las opiniones (más de 50000 tweets) en torno al tema electoral y por ello los gráficos se realizaron con aquellos pares de palabras que aparecieron 500 veces o más para captar los temas más relevantes, generándose el gráfico de bigramas que se muestra a continuación.

**Gráfico 8**  
Bigrama para el periodo  
septiembre-diciembre de 2020



Fuente: los autores

En el gráfico 8 se puede ver que el tema que fue recurrente en todos los periodos anteriores asociado al discurso del presidente de la Asamblea Nacional Juan Guaidó ahora desaparece de la palestra, y en su lugar aparecen otros contenidos:

- La Asamblea Nacional, con motivo de las próximas elecciones para renovarla. Recuperarla, desde el punto de vista del gobierno (Es de notar que este tema fue el más mencionado y esto lo detecta el grafo, mostrando los puntos azules presentes en él con un tamaño mayor que los presentes en los otros temas).
- El comando de campaña Darío Vivas y la juramentación de sus candidatos y equipo.
- El dirigente político Henrique Capriles Radonski a propósito de sus gestiones para lograr la liberación de algunos presos políticos.
- El presidente Nicolás Maduro.
- Elecciones parlamentarias libres.
- El salario mínimo.

### 3.3. Actividad de los usuarios, tendencias y análisis de sentimientos

En este subcapítulo se hace un análisis a un nivel más profundo. En este caso se determinaron los usuarios más activos, las tendencias y el análisis de sentimientos expresados en promedio en el conjunto de tweets.

En un primer análisis se encontró que en la muestra (126375 tweets) había tweets de 37423 usuarios distintos. La participación de estos usuarios durante este periodo de dos años ocurrió de la siguiente manera:

- 3578 usuarios emitieron tweets en el periodo 1.

- 3680 usuarios emitieron tweets en el periodo 2.
- 3304 usuarios emitieron tweets en el periodo 3.
- 4436 usuarios emitieron tweets en el periodo 4.
- 14356 usuarios emitieron tweets en el periodo 5.
- 19921 usuarios emitieron tweets en el periodo 6.

Es interesante determinar, en base a la actividad de todos estos usuarios, cuáles de ellos fueron los más activos y cuáles tuvieron mayor impacto en la red social Twitter, siendo replicados sus tweets o siendo nombrados con frecuencia por otros usuarios de la red. En lo que sigue se muestran los resultados obtenidos.

### 3.3.1. Los más activos

En la siguiente tabla se presentan los usuarios que emitieron la mayor cantidad de tweets en cada periodo analizado.

**Tabla 7**  
Usuarios más activos

Periodo	Usuario	Tweets
ene-abr 2019	@vozvenezolana	107
may-ago 2019	@vozvenezolana	124
sep-dic 2019	@ReporteYa	279
ene-abr 2020	@AlvaroT66804156	154
may-ago 2020	@Leonardo4619	268
sep-dic 2020	@AlvaroT66804156	286

Fuente: Los autores

Puede observarse que los usuarios @vozvenezolana y @AlvaroT66804156 estuvieron muy activos en más de un periodo. Además destaca que el usuario @AlvaroT66804156 (Historiador según su perfil) fue quien emitió más tweets, superando a el usuario @ReporteYa (cuenta oficial de la red venezolana de periodismo ciudadano Reporte Ya).

### 3.3.2. Los más influyentes

Un usuario influyente no es solamente aquel que tiene muchos seguidores, sino también aquel que consigue que su mensajes sean retuiteados o citados una gran cantidad de veces. Es por esto que se contabilizaron, por periodo, la cantidad de retweets, citas y tweets, que obtuvieron cada uno de los usuarios en la muestra, y en función del total de retweets obtenidos, en cada periodo, se determinaron los más influyentes del grupo de tuiteros en estudio. Los resultados obtenidos se muestran en la tabla 8.

**Tabla 8**  
Usuarios más influyentes

Periodo	Usuario	Tweets	Retweets	Menciones
ene-abr 2019	@eleonorabruzual	36	4931	14
may-ago 2019	@ReporteYa	74	1616	0
sep-dic 2019	@ReporteYa	279	2327	0
ene-abr 2020	@osmarycnn	10	7961	4
may-ago 2020	@MarleneOrtizAD	135	2172	0
sep-dic 2020	@GeneralPueblo	34	2814	0

Fuente: Los autores

En la tabla 8 se puede ver, por ejemplo, que el usuario @ReporteYa destaca de manera significativa en 2019. Para el año 2020 los usuarios que destacan entre los más influyentes son todos personajes públicos, como el

usuario @osmarycnn (periodista corresponsal de CNN) en el primer periodo de 2020, @MarleneOrtizAD (exalcalde, exdiputada y conductora de un programa de TV ) en el segundo periodo de 2020 y @GeneralPueblo (diputado) en el último periodo del 2020.

### 3.3.3. ¿De quién se habló?

Resulta interesante conocer cuáles usuarios son los que reciben más mención cuando se habla del tema electoral. Por ello nos enfocamos en identificar los usuarios más citados en cada uno de los seis periodos. Los usuarios más citados se muestran en la siguiente tabla.

**Tabla 9**  
Usuarios más citados

Periodo	Usuario	Menciones
ene-abr 2019	@jguaido	1390
may-ago 2019	@jguaido	996
sep-dic 2019	@jguaido	548
ene-abr 2020	@jguaido	1421
may-ago 2020	@nicolasmaduro	1045
sep-dic 2020	@nicolasmaduro	7174

Fuente: Los autores

En la tabla 9 claramente destaca, de enero de 2019 a abril de 2020, el usuario @jguaido (cuenta oficial del presidente de la Asamblea Nacional, Juan Guaidó). Mientras que para el cierre del año 2020, el usuario más citado fue @nicolasmaduro (cuenta oficial del presidente de la República, Nicolás Maduro) superando incluso la suma de las menciones del usuario @jguaido.

### 3.3.4. ¿Qué fue tendencia en este grupo de tweets analizado?

Los hashtags suelen indicar qué tópico trataron de poner en tendencia los usuarios al hablar del tema electoral, pero también suelen mostrar qué tendencias promovieron los usuarios para hacer visibles sus opiniones. En la siguiente tabla se muestra el *hashtag* más usado en cada uno de los seis periodos analizados.

**Tabla 10**  
Tendencias destacadas

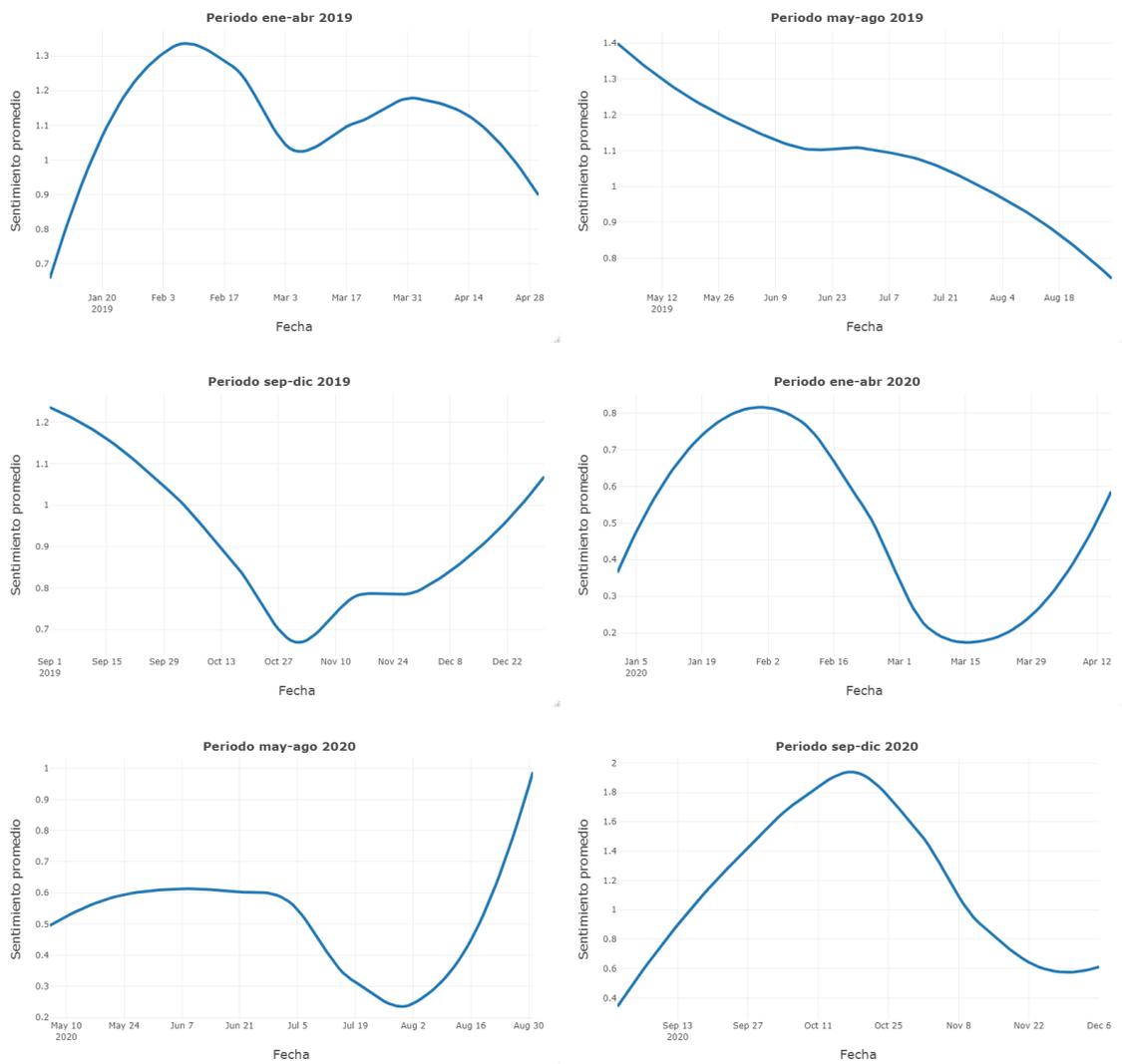
Periodo	Hashtag	Menciones
ene-abr 2019	#venezuela	187
may-ago 2019	#video	154
sep-dic 2019	#bolivia	181
ene-abr 2020	#cne	357
may-ago 2020	#cne	429
sep-dic 2020	#6d	3168

Fuente: Los autores

### 3.3.5. Análisis de sentimientos

Para dar una idea global de los sentimientos expresados por los usuarios, al hablar del tema electoral, se realizó por periodo un análisis de regresión. A continuación, en el gráfico 9, puede verse la evolución diaria del sentimiento en cada uno de los periodos.

**Gráfico 9**  
Sentimiento diario por periodo

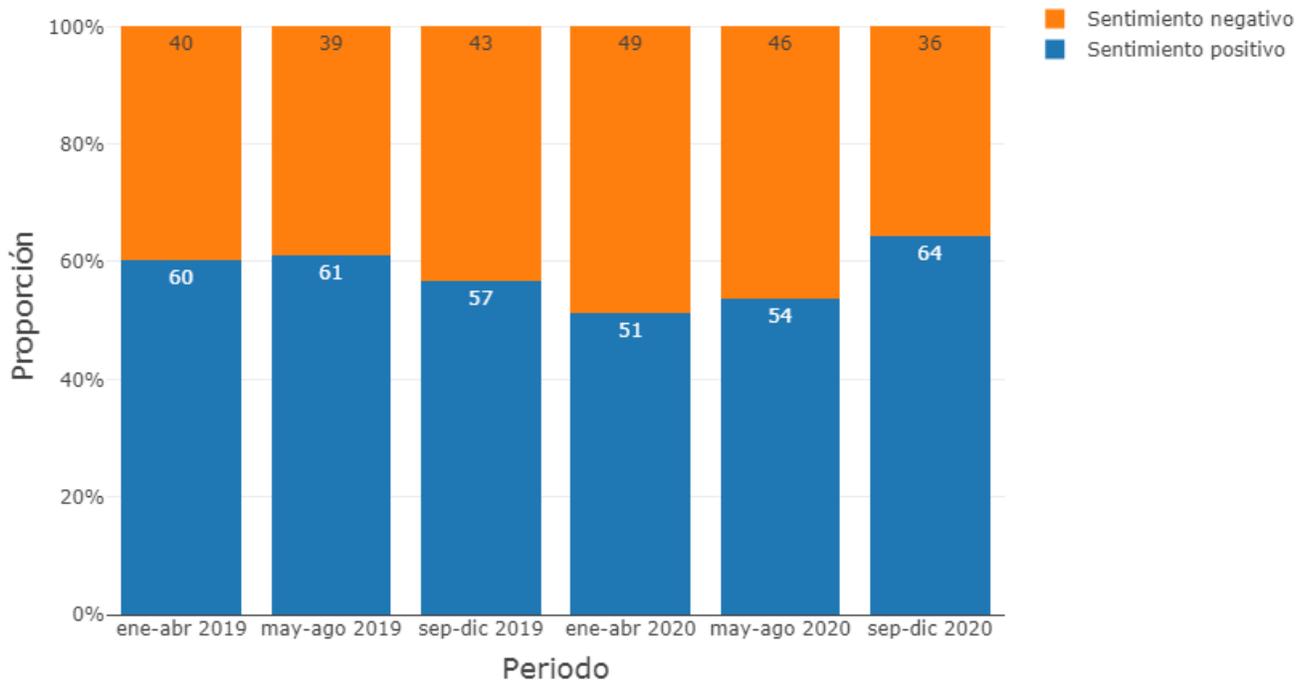


Fuente: Los autores

Como se puede apreciar en el gráfico 9, para el año 2019 el sentimiento, en promedio, es positivo. Se observa que, aún cuando es oscilante, con ascensos y descensos, siempre se mantiene por encima de 0.5 y alcanza su mayor valoración promedio en mayo de 2019. Para el año 2020, en el periodo enero-abril, la valoración promedio del sentimiento no logra llegar a 1 en la escala, obteniendo incluso valoraciones por debajo de 0.5, y no es sino hasta septiembre de 2020 que nuevamente alcanza una valoración promedio superior a 1, la cual se mantiene hasta noviembre de ese año, cuando nuevamente inicia su descenso.

Para finalizar, el siguiente gráfico muestra la proporción de sentimientos positivos y negativos que mostraron los tweets analizados por periodo.

**Gráfico 10**  
Proporción de sentimientos por periodo



Fuente: Los autores

## 4. Conclusiones

De acuerdo al análisis realizado para el año 2019 los venezolanos tuiteaban en los días de mayor actividad un número superior a los 400 tweets. Ya para el año 2020 los usuarios fueron mucho más activos y por esa razón alcanzamos a observar en los días de mayor actividad, hasta 2330 tweets.

Se evidenció que lo que impulsa las opiniones en esta plataforma son los temas políticos. Destaca que durante todo el año 2019 y hasta el primer periodo del 2020 el discurso del presidente de la Asamblea Nacional Juan Guaidó: cese a la usurpación, gobierno de transición y elecciones libres, se mantuvo en la palestra. De la misma forma resalta que el tema que generó más impacto fue el decreto presidencial del 31 de agosto de 2020, donde se otorga el indulto a más de cien presos políticos. Además resulta notorio como el tema de las elecciones en Bolivia movió la opinión de los venezolanos sobre el tema electoral y no perdieron la oportunidad de hacer comparaciones del evento electoral boliviano con los procesos electorales y la política en Venezuela.

Los personajes que obtuvieron mayor mención en la plataforma, al opinar sobre el tema electoral en Venezuela, fueron el presidente de la Asamblea Nacional Juan Guaidó y el presidente de la república, Nicolás Maduro.

Si bien es cierto en todos los periodos se observa un sentimiento, en promedio, positivo también destaca que éste mayormente se encuentra por debajo de 1. Considerando que la escala va de -5 a 5, entonces podemos afirmar que las opiniones son levemente positivas.

Por otro lado, puede concluirse, que la opinión negativa alcanzó sus mayores niveles en el periodo enero-abril de 2020.

Este trabajo muestra como la red social es rica en información y como la implementación de las técnicas de minería de datos y minería de texto permiten analizar y resumir las opiniones de los usuarios de esta plataforma. Como próximo objetivo nos planteamos clasificar los tweets y determinar si en fechas importantes, en cuanto a procesos electorales en Venezuela se refiere, es posible identificar la presencia de cuentas bots.

---

## Referencias bibliográficas

- Gualda, E. y Rebollo, C. (2020). Big Data y Twitter para el estudio de procesos migratorios: Métodos, técnicas de investigación y software. *EMPIRIA. Revista de Metodología de Ciencias Sociales*, (46), 147-177.
- Jo T. (2019). *Text mining: Concepts, Implementation, and Big Data Challenge*. New York: Springer.
- Pang, B. y Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, (2), 1-135.
- Silge, J. y Robinson, D. (2017). *Text Mining with R: A Tidy Approach*. Sebastopol, U.S.A.: O'Reilly Media.
- Suau-Gomila, G., Percastre-Mendizabal, S., Palà, G., y Pont, C. (2017). Análisis de la comunicación de emergencias en Twitter. El caso del Ébola en España. En: Sierra, J. & Liberal, S. *Uso y Aplicación de las Redes Sociales en el Mundo Audiovisual y Publicitario*. (119-130). Madrid, España: McGrawhill Education. ISBN: 9788448613570.

Esta obra está bajo una Licencia Creative Commons  
Atribución-NoCommercial 4.0 International

