

Cuantificando opiniones expresadas en tweets durante las Elecciones Presidenciales Ecuatorianas del 2021 por medio del Análisis de Sentimiento



Quantifying opinions expressed in tweets during the 2021 Ecuadorian Presidential Elections through Sentiment Analysis

López-Fierro, Saríah; Pacheco-Villamar, Rubén

Saríah López-Fierro

sariah.lopez@sansano.usm.cl

Universidad T. F. Santa María, Chile

Rubén Pacheco-Villamar

rpachecov@uees.edu.ec

Universidad Espíritu Santo UEES, Ecuador

Ecuadorian Science Journal

GDEON, Ecuador

ISSN-e: 2602-8077

Periodicidad: Semestral

vol. 5, núm. Esp.3, 2021

esj@gdeon.org

Recepción: 31 Agosto 2021

Aprobación: 04 Octubre 2021

URL: <http://portal.amelica.org/ameli/jatsRepo/606/6062738019/index.html>

DOI: <https://doi.org/10.46480/esj.5.3.156>

Los autores mantienen los derechos sobre los artículos y por tanto son libres de compartir, copiar, distribuir, ejecutar y comunicar públicamente la obra sus sitios web personales o en depósitos institucionales, después de su publicación en esta revista, siempre y cuando proporcionen información bibliográfica que acredite su publicación en esta revista. Licencia de Creative Commons Las obras están bajo una <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>



Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-SinDerivar 4.0 Internacional.

Como citar : López Fierro, S., & Pacheco Villamar, R. (2021). Cuantificando opiniones expresadas en tweets durante las Elecciones Presidenciales Ecuatorianas del 2021 por medio del Análisis de Sentimiento. Ecuadorian Science Journal, 5(3), 209-219. DOI: <https://doi.org/10.46480/esj.5.3.156>

Resumen: El presente trabajo resume los resultados obtenidos después de aplicar la técnica de Análisis de Sentimientos en tweets publicados durante el Debate y las Vueltas Presidenciales Ecuatorianas del 2021. Se exponen criterios y alternativas para organizar y cuantificar las opiniones expresadas en Twitter por miles de usuarios sobre los temas en mención, y los personajes involucrados. Estas técnicas podrían beneficiar a entes gubernamentales que buscan medir la aceptación ciudadana, u organismos privados que procuran conocer la opinión general sobre un servicio o una marca.

Palabras clave: Análisis de Sentimiento, Análisis de Datos, Twitter, Elecciones Presidenciales.

Abstract: This work summarizes the results obtained after applying the Sentiment Analysis technique in tweets published during the Presidential Debate and the Ecuadorian Presidential Rounds of 2021. Criteria and alternatives are exposed to organize and quantify the opinions expressed on Twitter by thousands of users on the issues in question, and the characters involved. These techniques could benefit government entities that seek to measure citizen acceptance, or private organizations that seek to know the general opinion about a service or a brand.

Keywords: Sentiment Analysis, Data Analysis, Twitter, Presidential Elections.

INTRODUCCIÓN

Debido a que, en las Redes Sociales, los individuos pueden acceder a información rápida, redundante y de forma selectiva (Garrett, 2006), es común encontrar en ellas, expresiones a favor o en contra de personajes públicos, organismos, ideas u otros usuarios,

Siendo Twitter, una plataforma que comúnmente es usada para “publicar y compartir” (Scherman, et. al., 2014) opiniones. La cual admite a sus participantes canalizar los datos masivos por la disponibilidad de “consentir que les sigan” (followers) o “el seguir” (following) (Fábrega, et. al., 2013) cuentas de sus preferencias; da paso a que sus usuarios accedan a una “relación” o comunicación inadvertida (Scherman, et. al., 2014), lo que favorece a escenarios de intercambios de ideas entre desconocidos.

Por lo tanto, dentro del contexto político, guerras de opiniones o contiendas invaden a esta red social, generando datos masivos, difíciles de controlar por individuos, pero factibles de interpretar por Analistas de Datos con el apoyo de tecnologías avanzadas. Quienes buscan cuantificar, con diferentes técnicas, las opiniones expresadas de forma pública.

Este artículo presenta los resultados obtenidos después de analizar tweets descargados durante dos eventos políticos ecuatorianos: el debate y las vueltas presidenciales ecuatorianas del 2021. Se incluyen los criterios de selección y descarte usados para poder obtener los resultados finales, los mismos que pueden ser aplicados en los diferentes sistemas de información que analizan datos.

El objetivo del presente documento es exponer criterios de Análisis de Datos que permiten cuantificar opiniones emitidas en las Redes Sociales, y que podrían contribuir para la toma de decisiones en organismos gubernamentales, entidades privadas, personajes públicos, etc.

DATOS Y METODOLOGÍA

Datos

Por medio del API de Twitter, se descargaron los tweets referentes al debate presidencial y a las dos rondas de elecciones presidenciales del 2021.

Debate Presidencial

- En el primer evento se usó como filtro para descargar tweets, las cuentas de usuarios de todos los candidatos a la presidencia y la etiqueta #DebatePresidencialEc.

En total se descargaron 487.873 tweets, y después de su procesamiento, se obtuvieron 69.451 publicaciones únicas, las que fueron analizadas

Vueltas Electorales

- Durante el periodo de las vueltas electorales presidenciales, solo se descargaron tweets que incluyeran en su contenido las cuentas de los candidatos finalistas, Andrés Arauz y Guillermo Lasso, además de sus apellidos; y otras referencias en las que se hacía mención de ellos.

Así también, través de la API de Twitter se descargaron 1'302.388 tuits durante el periodo del 3 de febrero al 27 de abril, tiempo que comprendió las dos vueltas electorales.

Esta primera base de datos, sin procesar, consistía en tweets únicos, tweets duplicados y retweets. Por ello, como primer paso, separamos los retweets en una base de datos diferente y descartamos los tweets duplicados. Lo que nos dio un total de 298,326 tweets únicos. El análisis de sentimiento se aplicó a este último conjunto. No obstante, el número de retweets realizados por los usuarios también se sintetizó y se los asignó a cada uno respectivamente, en el resumen final.

Metodología

La principal técnica usada fue el Análisis de Sentimiento de los tweets. A partir de este paso, procesamos otras vistas o presentaciones de resultados factibles debido a los metadatos de cada tweet descargado. A continuación, se explican las metodologías aplicadas para obtener dichos resultados.

Análisis de Sentimiento

- El Análisis de Sentimiento busca automáticamente monitorear y medir a tiempo, los comentarios u opiniones de los usuarios y, marcar el tono emocional encontrado en éstos (Feldman, R., 2013).

Nuestro enfoque usado en el Análisis de Sentimiento fue por la comparación de palabras con diccionarios.

Para ello, elaboramos dos diccionarios (uno positivo y otro negativo) de palabras, “emojis” y “hashtags” (o etiquetas) que incluían expresiones o jergas ecuatorianas. Además, se incluyeron conjugaciones (correspondientes al idioma español) en el caso de los verbos, y las variaciones entre géneros y números que podrían aplicar en todos los adjetivos incluidos. Estos diccionarios fueron actualizados constantemente, al encontrar en nuevos tweets diferentes palabras o expresiones informales pero que eran comunes en el ambiente político que se examinaba.

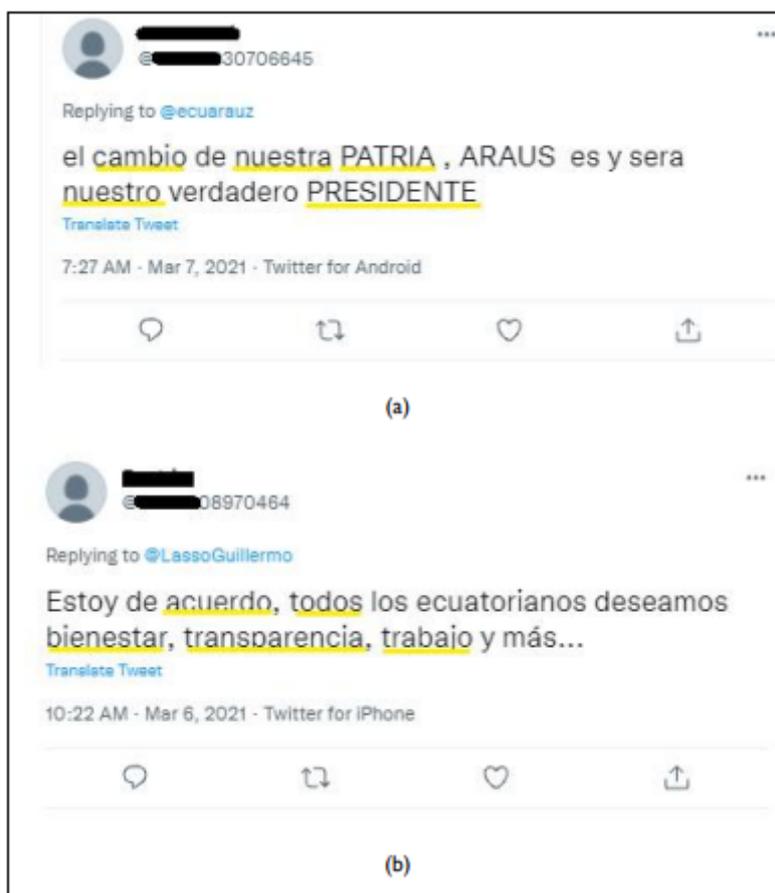


FIGURA 1

Ejemplos de tweets dirigidos (a) Andrés Arauz y (b) Guillermo Lasso con polaridad positiva.

Adicional a ello, se calificó cada palabra de los diccionarios del 1 al 3, dando 1, comúnmente a aquellas palabras negativas/positivas generales (como por ejemplo del diccionario negativo: Apenaron, Arruinó, Ausente, etc.). Asignábamos el valor de 2, comúnmente a adjetivos calificativos positivos/negativos

(como por ejemplo del diccionario negativo: Baboso, Borracho, Ignorante, etc.). Y la calificación de 3, comúnmente a etiquetas que inequívocamente expresaban una posición o polarizaban el “sentimiento” del tweet (como por ejemplo del diccionario negativo: #AndrésNoMientasOtraVez, #LassoEsMoreno, #ElPeorGobiernodelaHistoria, etc).

Luego, comparamos cada palabra del tweet, con cada palabra de los diccionarios. Si existía un “match” o se detectaba la presencia de la palabra, se ponía la calificación correspondiente de ésta (a las palabras positivas se las marcaba con polaridad positiva, mientras que a las palabras negativas se las marcaba con polaridad negativa). Una vez que todas las palabras estaban calificadas, se “sumaban”. Si el resultado final era positivo, se calificaba al tweet, como tal. Si era negativo, con esta misma polaridad; si era “0” su resultado, se lo marcaba como tweet “Neutro”; y finalmente, si ninguna palabra era encontrada en los diccionarios, recibía la etiqueta de “Sin Calificar”.

Por lo tanto, todos los tweets únicos fueron etiquetados con estos cuatro estados: “Positivo”, “Negativo”, “Neutro” y “Sin Calificar”.

Además, el resultado de la suma final era promediado con el total de palabras encontradas en el tweet, lo que exponía en efecto un porcentaje de polarización. En consecuencia, entre mayor era el porcentaje de los tweets marcados como “positivos”, mayor era su probabilidad de acierto. De la misma forma sucedía con los tweets marcados como negativos.

Por ejemplo, en la Figura 1 se presentan dos ejemplos de tweets que fueron marcados como positivos, y tuvieron el mismo porcentaje de polaridad. Las palabras subrayadas de amarillo fueron detectadas por el diccionario de palabras positivas.

Asignación de Polaridad a los candidatos

- Una vez conocida la polarización de los tweets, un dato de interés es conocer hacia quién está dirigido cada tweet. Previamente se mencionó que los filtros de descarga incluyeron las cuentas oficiales de los candidatos (@ecuarauz y @lassoguillermo), así también como sus apellidos (Arauz y Lasso) y otras palabras que se popularizaron en el contexto político ecuatoriano para hacer referencias a ellos (por ejemplo, sánduche, banquero, etc).

No obstante, estos filtros descargaron también tweets que contenían, por ejemplo, los apellidos de los candidatos, pero que hacían referencia a otro personaje o evento. Por ejemplo, un considerable grupo de tweets que fueron descargados porque incluían el apellido “Lasso”, eran del programa estadounidense “Ted Lasso”. Una alternativa que encontramos para evitar este tipo de falsos-positivos, fue descartar los tweets en los que el lenguaje que se marcaba en los metadatos no era español.

Así también, notamos que, aunque un tweet mencionaba a uno de los candidatos y estaba marcado como negativo, no siempre sus palabras eran dirigidas a él. En algunos casos, eran sus respuestas a otras cuentas. Para evitar errores en la asignación de polaridad, detectamos el número de cuentas a las que un Tweet hacía mención, y filtramos solo a “1” (@ecuarauz o @lassoguillermo), esta nueva consulta.

Luego de esta acción, separamos en nuevas bases de datos, todos los tweets que fueron dirigidos solo hacia Andrés Arauz y en otra a Guillermo Lasso, con su respectiva polarización y demás metadatos. Este proceso nos permitió analizar las palabras usadas en relación con el apoyo o crítica hacia cada uno.

Geolocalización

- Twitter permite a sus usuarios registrar voluntariamente la ubicación. Debido a que este, no es un campo verificado, cualquier localidad escrita de cualquier forma, puede ser colocada en este espacio.

Por ejemplo, al analizar las localidades de los tweets descargados, nos encontramos con las siguientes maneras de hacer referencia a la ciudad de Guayaquil: “Guayaquil city”, Gye, “Perla del

Pacífico”, GuayaKill, “en el Monumental”, “Ciudad del Río Grande, Guayas”, “Guayaco”, “Lat 0° - Guayaquil”, “ ☆Guayaquil☆ ↔ ★Ecuador★ ”, entre otros. Tal como, distintas formas de hacer referencia a la ciudad de Quito: “uio”, “Ecuador – Quito”, Kito/Ecuador, “LDU Quito”, “en la mitad del mundo”, Quitof, Quitoff, Quitofff, entre otros.

Por ende, con el fin de cuantificar las opiniones de los usuarios de Twitter de acuerdo con su localidad, buscamos dentro de una muestra de tweets, las diferentes formas en las que los usuarios describían sus ciudades, las organizamos a nivel de provincias; y elaboramos un nuevo diccionario con esta nueva relación.

De esta forma en todo usuario cuya localidad se encontraba escrito “Gye”, por ejemplo, automáticamente era cambiado a “Ecuador, Guayas”. O en casos de ubicaciones internacionales, si se detectaba registrado “Madrid”, por ejemplo, era automáticamente cambiado a “España”.

Luego de realizar esta estandarización, automatizamos la asignación de coordenadas geográficas en todos los usuarios de acuerdo con su ubicación, y distinguimos en los mapas creados todas las incidencias positivas o negativas según el candidato mencionado.

También incluimos localidades internacionales, y las organizamos a nivel de países, solo para fines de contabilización. Pero a para representar la polaridad en un mapa geográfico, permitimos que la asignación de coordenadas sea hecha conforme a la localidad dada por los usuarios. En la Figura 2, se muestra un ejemplo de la distribución de los tweets dirigidos hacia los respectivos candidatos según su polarización (banderín amarillo representan los tweets positivos, y banderín rojo representan los tweets negativos) y ubicación.

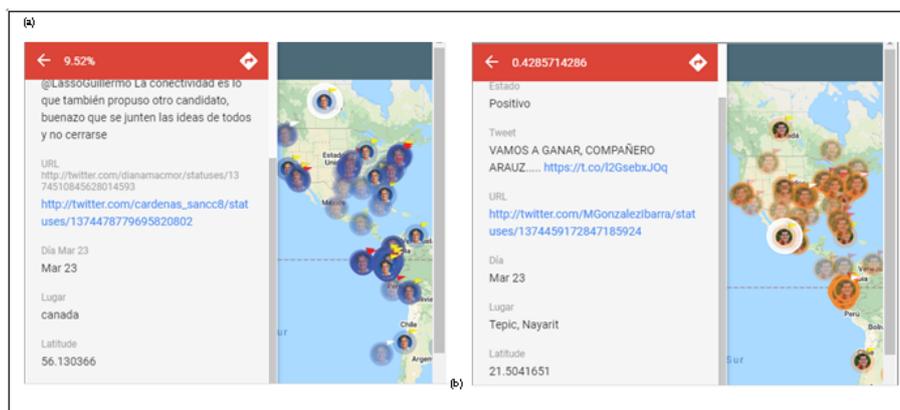


FIGURA 2.
Distribución de tweets dirigidos hacia (a) Andrés Arauz y (b) Guillermo Lasso con polaridad positiva, y colocados de según su geolocalización.

Etiquetas y palabras principales

- Otro modo de analizar datos masivos es por medio de la cuantificación de palabras.

Para poder obtener algún dato importante en este proceso, eliminamos primero dentro de todos los tweets, “artículos”, “adverbios”, etc., o cualquier subclase de palabra que sirva como complemento de aquellas que representan las ideas principales.

En este paso, no sólo destacamos las palabras más comunes usadas al referirse a un candidato en el debate presidencial. Si no también las etiquetas que tuvieron mayor propagación en la red.

Este es el primer paso para un análisis más detallado sobre el análisis de palabras. Técnicas de NPL, podrían adaptarse también para encontrar las relaciones de éstas, crear clústeres con las mismas, organizarlas por su información léxica o sinónimos bajo un contexto específico, entre otros (Aouicha,

M. B., et. al., 2016). Y de esta manera, conocer, por ejemplo, cuáles son los aspectos negativos que más destacan los usuarios sobre un candidato, y con cuáles eventos u otros personajes lo relacionan.

Etiquetas y palabras principales

- Otro modo de analizar datos masivos es por medio de la cuantificación de palabras.
Para poder obtener algún dato importante en este proceso, eliminamos primero dentro de todos los tweets, “artículos”, “adverbios”, etc., o cualquier subclase de palabra que sirva como complemento de aquellas que representan las ideas principales.
En este paso, no sólo destacamos las palabras más comunes usadas al referirse a un candidato en el debate presidencial. Si no también las etiquetas que tuvieron mayor propagación en la red.
Este es el primer paso para un análisis más detallado sobre el análisis de palabras. Técnicas de NPL, podrían adaptarse también para encontrar las relaciones de éstas, crear clústeres con las mismas, organizarlas por su información léxica o sinónimos bajo un contexto específico, entre otros (Aouicha, M. B., et. al., 2016). Y de esta manera, conocer, por ejemplo, cuáles son los aspectos negativos que más destacan los usuarios sobre un candidato, y con cuáles eventos u otros personajes lo relacionan.

Trolls de internet

- Aunque es común que, durante el tiempo de campañas, la temática de las redes se torne a publicar información a favor o en contra de algún candidato. Sin embargo, es sospechoso encontrar una elevada participación política desde cuentas que tienen poco tiempo de creación y que divulgan elevados mensajes similares, múltiples, duplicados y polarizados sobre el evento político a acontecer. En estos casos podríamos encontrarnos ante los llamados “trolls”.
Uno de los propósitos de los trolls de internet es provocar argumentos entre personas, y así manipular la percepción pública (Bradshaw et. al., 2017; Glisson, 2019), a favor o en contra de algún personaje, idea o evento.
Los siguientes procesos fueron realizados para destacar cuentas sospechosas.
Una vez aplicado el Análisis de Sentimiento en los tweets, los organizamos conforme a los usuarios de Twitter que los crearon. Luego resumimos por usuario, la cantidad de tweets positivos y negativos que habían enviado hacia Andrés Arauz y/o Guillermo Lasso. Así también, sumamos el total de retweets y, distinguimos cuando habían compartido contenido de Andrés Arauz o, de Guillermo Lasso.
Con estos datos, notamos que un grupo considerable de cuentas que recientemente habían sido creadas, tenían una importante actividad e inclinación a favor o en contra de uno de los candidatos. En consecuencia, con el fin de “encontrar trolls”, otorgamos una nueva calificación por cada una de las siguientes características que era detectada en los usuarios: (1) poco tiempo de creación (a partir del 2019), (2) conserva el mismo nombre de usuario que automáticamente otorga Twitter, (3) número de retweets superior a los mensajes publicados, y (4) número de “followers” mayor que el número de “following”.
Estos criterios expusieron varias cuentas de Twitter. En la Figura 1, se exponen los tweets de dos cuentas que fueron marcadas como “Trolls” al cumplir con las características mencionadas.
Contabilizamos por usuarios los valores dados por cada característica cumplida, y obtuvimos como dato final el número de cuentas que podrían estar “manipulando la percepción pública”.



FIGURA 3

Nube de palabras de tweets negativos hacia (a) Andrés Arauz y (b) Guillermo Lasso.

RESULTADOS Y DISCUSIONES

Esta sección incluye resultados, de acuerdo con las metodologías y procesos previamente descritos. Cabe recalcar que existen diversas alternativas en las que los datos podrían adaptarse, filtrarse y mostrarse, mas con los ejemplos incluidos se responde ante los criterios descritos en la sección anterior.

Debate Presidencial

Del total de tweets originales, el 15,39% correspondieron a publicaciones dirigidas solo a @lassoguillermo o lasso. Mientras que el 10,29% fueron dirigidas solo a @ecuarauz o arauz.

Las 5 etiquetas más usadas fueron #debatepresidencialec, #lassonoresponde, #lassoesmoreno, #ecuador, #andresnomientasotravez. Las 5 cuentas que mayor mención recibieron fueron: @lassoguillermo, @ecuarauz, @mashirafael, @cnegobec, @rabascallcarlos.

Luego de procesar el sentimiento de todos los tweets, organizarlos según el candidato al que se estaban dirigiendo, contabilizamos en una sección distinta solo las palabras de los tweets que hacían mención negativa hacia Andrés Arauz y, en otra, todas las palabras de los tweets que hacían mención negativa hacia Guillermo Lasso, y los representamos en una “nube de palabras” (gráfico adecuado para exponer las palabras que tienen mayor prominencia en un texto, debido a que destaca con mayor tamaño las palabras que más veces se encontraron en el texto o los datos computarizados).

Después de eliminar las menciones a dichos candidatos y la etiqueta #debatepresidencialec, debido a que eran las que mayor número de incidencias tenían, se generaron los resultados presentados en la Figura 3. Es posible notar que en el gráfico (a), los tweets negativos dirigidos a Andrés Arauz incluyeron en su mayoría las palabras “país”, “correa”, “todo”. Mientras que los tweets negativos dirigidos hacia Guillermo Lasso, tenían varias predominantes, pero en un nivel ligeramente superior se encontraron: “#lassoesmoreno”, “#lassonoresponde”, “pueblo”.

Así como fue explicado previamente, un análisis consecuente a realizar con estos resultados es la clusterización de palabras, que permitiría entender qué otros factores interconectan el malestar expresado de las personas hacia Andrés Arauz y su relación con Correa. O en el caso de Guillermo Lasso, qué otras

palabras comúnmente van acompañadas de la etiqueta que lo relacionan a Lenín Moreno, y por lo tanto el uso constante de la etiqueta.

Rondas Electorales

Una vez filtrados los tweets duplicados, separados los retweets, y aplicado en los tweets restantes el Análisis de Sentimiento, se obtuvieron los resultados presentados en la Tabla 1,

TABLA 1.
Número total de Tweets positivos y negativos publicados durante las Elecciones Presidenciales

Tabla 1. Número total de Tweets positivos y negativos publicados durante las Elecciones Presidenciales

Total	Positivo	Negativo	Neutros y Sin Calificar
298.326	97.819	118.327	57.910

Con los tweets polarizados es posible obtener distintas vistas de resultados. Entre ellas, para conseguir un porcentaje elevado de asertividad, procedimos a filtrar los tweets que sólo mencionaban a los candidatos finalistas, descartando de esta forma, toda mención a otro usuario. La Tabla 2, muestra los resultados de este proceso.

Entre otras vistas de resultados, luego de estandarizar las ubicaciones compartidas por los usuarios de nuestra Base de Datos a nivel provincial, cuando especificaban ciudades del Ecuador, obtuvimos los resultados expuestos en la Tabla 3. Nótese que algunas provincias muestran una aprobación contraria a las publicadas en los resultados finales, como por ejemplo, Guayas. En esta provincia, Andrés Arauz ganó oficialmente con un 53.05%, mientras que en nuestros resultados, se aprecia que de todos los tweets procesados y polarizados, en los que se mencionó a Andrés Arauz, el 58.15% publicó tweets negativos. Mientras que del total de menciones polarizadas dirigidas solo a Guillermo Lasso, el 39.47% fueron negativas, y por lo tanto un 60.52% positivas. Dando como ganador a Guillermo Lasso.

TABLA 2.
Número de tweets positivos y negativos dirigidos sólo hacia Andrés Arauz o Guillermo Lasso.

Arauz		Lasso	
Positivo	Negativo	Positivo	Negativo
3.513	8.316	9.800	14.494
Total: 13.652		Total: 29.177	

A nivel internacional, se encontraron tweets provenientes de los cinco continentes. La Figura 4, destaca por los colores de campaña correspondientes, los países en las que fueron ganadores los candidatos.

Pueden darse algunas explicaciones a estos resultados. Una de ellas es que solo 1.1 millones de ecuatorianos con acceso a Redes Sociales, usan Twitter[3]. Es decir que, nuestra base de datos es una muestra del 7.69% de ecuatorianos, y reduciendo aún más el grupo a solo aquellas cuentas que publicaron su ubicación en su perfil. Otra razón, podría justificarse en que los resultados se los presenta a nivel de provincia y no por ciudad. En

los resultados oficiales, Guillermo Lasso tuvo mayor aceptación en Guayaquil, mas perdió a nivel provincial por las otras ciudades.

A nivel internacional, se encontraron tweets provenientes de los cinco continentes. La Figura 4, destaca por los colores de campaña correspondientes, los países en las que fueron ganadores los candidatos.



FIGURA 4
Mapa geográfico que representa los países en las que los candidatos, de acuerdo con los colores de campaña, tuvieron mayores opiniones positivas.

Finalmente, otro resultado obtenido después de Analizar el Sentimiento fue el encontrar cuentas en la base de datos que comúnmente publicaban tweets con porcentajes altos de polarización a favor o en contra de uno de los candidatos.

Con la finalidad de hallar patrones en estas cuentas, las filtramos de acuerdo con los criterios descritos en la sección anterior: (1) la cantidad de número de retweets realizados por usuarios, (2) la comparación entre el número de “followers” inferior al número de “following”, y (3) el uso de nombre de cuenta entre los usuarios, predefinido por Twitter, y los distribuimos por el año de creación.

TABLA 3.
 Número de tweets positivos y negativos dirigidos sólo hacia Andrés Arauz o Guillermo Lasso de acuerdo con las provincias.

	Arauz		Lasso	
	Positivo	Negativo	Positivo	Negativo
Ecuador	928	983	1.816	1.308
Azuay	99	114	249	173
Bolívar	9	3	19	4
Cañar	8	7	18	2
Carchi	2	3	4	1
Chimborazo	13	26	47	30
Cotopaxi	3	23	31	15
El_Oro	48	43	111	86
Esmeraldas	27	17	53	21
Galápagos	13	0	9	16
Guayas	657	913	1.811	1.181
Imbabura	22	33	23	29
Loja	49	48	83	47
Los_Rios	19	13	40	23
Manabi	123	94	254	175
Morona_Santiago	3	2	18	5
Napo	6	1	13	8
Orellana	4	9	17	7
Pastaza	0	4	1	2
Pichincha	718	1.202	2.008	1.425
Santa_Elena	10	17	22	28
Santo_Domingo_Tsachilas	17	14	45	46
Sucumbios	9	5	14	12
Tungurahua	30	44	61	41
Zamora_Chinchi	3	2	3	3

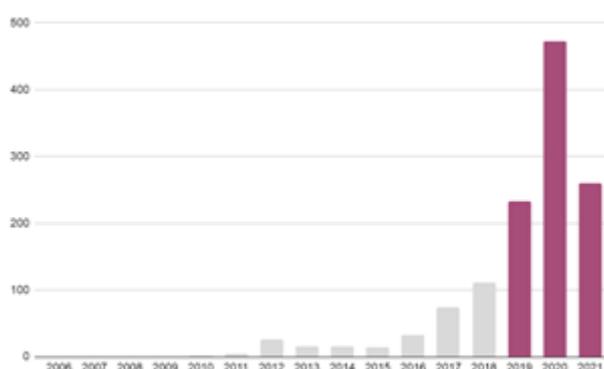


FIGURA 5.
 Número de trolls distribuidos según el año de creación.

La Figura 5 presenta el total de cuentas que fueron marcadas como “trolls” de acuerdo con el año de creación. Nótese que el mayor número de cuentas fueron recientemente creadas en los años 2019, 2020 y 2021.

Se encontró evidencia de cuentas con características sospechosas que publicaban y compartían información para perjudicar a uno de los dos candidatos, y así también cuentas que buscaban apoyarlos. Por ello, podemos concluir que ambos candidatos se vieron beneficiados y perjudicados por los llamados trolls de internet.

CONCLUSIONES

Medir las opiniones expresadas en redes sociales, es conocer a un grupo de ciudadanos o consumidores.

En este artículo se han expuesto criterios y alternativas para organizar y cuantificar datos públicos facilitados por Twitter. Estas técnicas podrían beneficiar a entidades privadas o públicas que requieran de alternativas más eficientes para conocer la aceptación de una idea o producto, por ejemplo.

Como trabajo futuro se ha planificado analizar el nivel de influencia de un tweet, con el fin de encontrar patrones que podrían permitir medir el nivel de propagación de personajes públicos, “influencers”, cuentas activistas o trolls.

El ser parte de la generación del internet, también implica el ser consumidores de datos masivos. El conocer las técnicas más eficientes para poder procesarlos, es parte de los desafíos actuales.

AGRADECIMIENTOS

Mención especial a Carlos Chiriboga Calderón por contribuir en la realización de esta investigación.

Esta investigación ha sido apoyada y financiada por el Departamento de Investigaciones Tecnológicas de la empresa Soluciones Wandarina S. A. de Ecuador.

REFERENCIAS BIBLIOGRÁFICAS

- Aldunate, X. F. (2021). La propaganda política a través de la utilización de "fake news": elecciones Ecuador 2021, preparándonos para la desinformación. *Cosmovisión de la comunicación en redessociales en la era postdigital*, 229
- Aouicha, M. B., Taieb, M. A. H., & Hamadou, A. B. (2016). Taxonomy-based information content and wordnet-wiktionary-wikipedia glosses for semantic relatedness. *Applied Intelligence*, 45(2), 475-511.
- Garrett, K. R. (2006). Protest in an information society: A review of literature on social movements and new ICTs. *Information, communication & society*, 9(02), 202-224.
- Scherman, A., Arriagada, A., & Valenzuela, S. (2014). Student and Environmental Protests in Chile: The Role of Social Media. *Politics*, 35(2), 151–171. doi:10.1111/1467-9256.12072
- Fábrega, J., & Paredes, P. (2013). La política chilena en 140 caracteres. *Intermedios. Medios de comunicación y democracia en Chile*, 199-224.
- Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4), 82-89.
- Bradshaw, S., & Howard, P. (2017). Troops, trolls and troublemakers: A global inventory of organized social media manipulation.
- Glisson, L. (2019). Breaking the spin cycle: Teaching complexity in the age of fake news. *Libraries and the Academy*, 19, 3, 461-484. <https://doi.org/10.1353/pla.2019.0027>