

Comunicación política en Twitter y su análisis automático mediante el uso de datos ordenados y visualización de información

Rocío Abascal Mena

Universidad Autónoma Metropolitana, Unidad Cuajimalpa,
Ciudad de México, México
mabascal@correo.cua.uam.mx

Resumen. Actualmente, las redes sociales cobran vital importancia debido a que plantean una nueva forma de comunicación a través de medios digitales, donde su alcance es local y global, y su tiempo de difusión es inmediato. Cada vez más, en los procesos electorales no es sorprendente una mayor presencia de los políticos para hacer campañas y ganar contiendas debido a su escasa regulación y a la alta potencialidad para evadir las restricciones que se ponen a las campañas anticipadas. Independientemente de las regulaciones, la presencia de movimientos ciudadanos en diversos ámbitos de la vida diaria muestra claramente que las redes sociodigitales crean una ola polifónica y, definitivamente, un gran gesto de democratización. Este artículo muestra los primeros resultados de un estudio de caso, a partir del uso de datos ordenados y la visualización de información, sobre los perfiles y tuits de los candidatos a Jefe de Gobierno de la CDMX en 2018. Se explican algunas limitaciones de las técnicas actuales para el análisis de tuits y finalmente se presentan unas breves conclusiones con el fin de presentar las posibilidades de un análisis automático que pueda realizado de manera rápida.

Palabras clave: datos ordenados, análisis de redes sociales, visualización de información, comunicación política, Twitter.

Political Communication on Twitter and its Automatic Analysis through the Use of Ordered Data and Information Visualization

Abstract. Currently, social networks are vitally important because they pose a new form of communication through digital media, where its reach is local and global, and its time of dissemination is immediate. Increasingly, in the electoral processes it is not surprising to see a greater presence of politicians to make their campaigns and win contests due to the lack of regulation and the high potential to evade the restrictions placed on the advance campaigns. Regardless of the regulations, the presence of citizen movements in various areas of our daily life clearly shows that socio-digital networks create a polyphonic wave and, definitely, a great gesture of democratization. This article shows the first results

of a case study, from the use of sorted data and the visualization of information, about the profiles and tweets of the candidates for Head of Government of the CDMX in 2018. Some limitations of the techniques for the analysis of tweets are explained and finally some brief conclusions are presented in order to show the possibilities of an automatic analysis that can be done in quickly way.

Keywords: ordered data, social network analysis, information visualization, political communication, Twitter.

1. Introducción

Actualmente, las redes sociales cobran vital importancia debido a que plantean una nueva forma de comunicación a través de medios digitales, donde su alcance es local y global, y su tiempo de difusión es inmediato. Una red social está definida como: “un conjunto organizado de personas formado por dos tipos de elementos: seres humanos y conexiones, entre ellos, las cuales evolucionan orgánicamente a partir de la tendencia natural de toda persona a establecer relaciones y hacer pocos o muchos amigos” [1].

En los últimos años, el uso de las redes sociodigitales ha desempeñado un papel cada vez más visible en la política. Por ejemplo, un número creciente de pakistaníes se dirigió a la gente a través de mensajes de texto, como medios de comunicación alternativos durante la Emergencia de Pakistán del 2007 al 2008. Esto comenzó después de que el presidente del país, el General Pervez Musharraf, suspendió al Jefe de Justicia de la Suprema Corte. En consecuencia, el gobierno canceló redes de teléfonos celulares y bloqueó los canales de noticias [2]. En el año 2010, en algunos países árabes, como Irán, se llevó a cabo la llamada “Primavera Árabe” a través de protestas a gran escala que fueron coordinadas a partir de Twitter.

En México, se recuerdan movimientos políticos generados inicialmente en las redes sociales como #Yosoy132 en el 2012, #Ayotzinapa, #TodosSomosIPN, #nosfaltan43 en el 2014 o #Gasolinazo2017 en 2017. Todos ellos tienen características semejantes como es el uso de las redes sociales para manifestar el repudio de hechos suscitados permitiendo llegar a cientos o miles de seguidores. Esto se debe en gran medida a una regla incontestable del juego: ni emisores ni receptores registran su verdadero nombre. De igual manera, en los procesos electorales no es de sorprenderse una mayor presencia de los políticos para hacer campañas y ganar contiendas debido a su escasa regulación y a la alta potencialidad para evadir las restricciones que se ponen a las campañas anticipadas.

Independientemente de las regulaciones, la presencia de movimientos ciudadanos, en diversos ámbitos de la vida diaria, muestra claramente que las redes sociodigitales crean una ola polifónica y, definitivamente, un gran gesto de democratización.

En las siguientes secciones se presenta: 1) el trabajo relacionado con el análisis de Twitter y la política desde la Computación, 2) se muestran los primeros resultados de un estudio de caso sobre los perfiles y tuits de los candidatos a Jefe de Gobierno de la CDMX en 2018 a partir del uso de datos ordenados, 3) se explican algunas limitaciones de las técnicas actuales para el análisis de tuits y 4) finalmente se presentan unas breves conclusiones.

2. El estudio de Twitter a través de la computación

El estudio de las redes sociales aunado a la política es un campo de investigación que resulta interesante e importante al aprovechar la información generada, por la sociedad, en la que las formas de interacción social están mediadas por el uso de las nuevas tecnologías. En Twitter, la continua evolución de las etiquetas, precedidas por el símbolo # y nombradas como hashtags, permite seguir el acontecimiento diario a partir de los ojos de miles de ciudadanos que son organizadores, y protagonistas creando símbolos que permiten la colaboración en red. Podemos, así, observar una red social que se ha convertido en un espacio para el humor, el ingenio y el discurso directo.

Twitter es una red social bastante efectiva al aprovechar en la mínima expresión de sus mensajes la posibilidad de llegar al otro. Esta característica, que nos remite a tiempos del uso del telegrama, resulta paradójica al contrastarla con el flujo incesante de información que tenemos hoy en día. La restricción en el uso de caracteres parece ser una ventaja que incita a una rápida publicación, recepción y visualización de los mensajes. Así, la principal ventaja es la facilidad e invitación a la participación masiva, como la creación de recursos encaminada a la optimización de la significación y el espacio.

En el área de la Computación existen numerosas investigaciones que recuperan información proveniente de Twitter ya sea a través de un determinado hashtag (#) o de un actor (@). A partir de los datos recuperados es posible analizar a los individuos y/o colectivos que componen los mensajes, determinar la posición de los actores en la red, conocer la intención del mensaje (si es positivo o negativo) o analizar la evolución de la red a través del tiempo, por citar algunos ejemplos de aplicación.

En relación a los métodos de análisis de datos obtenidos de Twitter, en [3] se describen los tres más recurrentes los cuales son: 1) análisis de contenido: estudio de tuits mediante técnicas de minería de texto; 2) análisis de redes: detección de patrones de relaciones entre usuarios y 3) análisis de sentimiento: clasificación de tuits de acuerdo con la actitud expresada en ellos. Sin embargo, una gran mayoría de las investigaciones que utilizan Twitter como su fuente principal de datos realizan lo que se llama análisis de sentimiento y minería de opiniones. Para ello se requieren técnicas provenientes del área de Procesamiento del Lenguaje Natural, la Lingüística computacional y la Minería de textos. Este último, también se refiere a la minería de opiniones y al análisis de sentimientos. Ambos están relacionados con la obtención de información a través del análisis de las emociones que experimenta el usuario o su punto de vista acerca de algo en particular (reseñas, *posts*, etc.). El cálculo de la inclinación promedio de la opinión hacia cualquier entidad no solo ayuda a las organizaciones comerciales a saber más sobre un posible cliente, en el caso de Twitter es un excelente medio para conocer el punto de vista político del público en general. A su vez, la minería de opiniones ayuda a las organizaciones políticas a tener un mejor entendimiento de sus contrincantes [4].

Algunos estudios sugieren que en el ámbito político y de movimientos sociales basados en redes sociodigitales, los mensajes que aumentan su ira son los más virales y predisponen a los manifestantes a ir más decididos a votar [5], [6]. En el mismo sentido, recientemente, está el trabajo de [4] que afirma que una considerable proporción del discurso político de las elecciones nacionales se lleva a cabo en las redes sociodigitales. A partir del estudio de los temas de mayor prevalencia durante un

determinado tema, se investigó la variación temporal de los sentimientos positivos y negativos a la vez que se examinó la proximidad semántica entre los distintos temas. En este caso, el estudio se realizó en el contexto del referéndum griego del 2015 y las subsecuentes elecciones legislativas en el que como resultado se obtuvo que existía un fuerte sentimiento de anti-austeridad, acompañado con una visión crítica sobre las acciones políticas europeas y griegas [4].

En general, en el contexto mexicano, la gran mayoría de los análisis de Twitter se han centrado en relaciones político-ciudadano, procesos decisorios y movimientos sociales [7,8,9]. Además de ello, es de interés el estudio de los *bots* como un elemento distractor en protestas y contiendas electores. Tal es el caso de investigación que se presenta en [10] y en la que se analizan a los *bots* durante el movimiento mexicano #YaMeCansé.

En la siguiente sección se presenta el uso de datos ordenados a partir de perfiles de Twitter con el fin de analizar tuits recuperados durante un periodo de tiempo.

3. Uso de datos ordenados para el análisis de tuits a partir de perfiles de usuario

Uno de los principales intereses en el estudio de los tuits tiene que ver con el seguimiento de las cuentas o perfiles de los usuarios (precedidos por el símbolo @) debido a que la sola recuperación de información puede darnos indicios acerca de la influencia o liderazgo de estos.

El año 2018, para México, se caracterizó por su intensa actividad comicial dado que se llevaron a cabo elecciones federales como locales en todos los estados del país, incluyendo la Ciudad de México (CDMX). Para la CDMX, en la contienda, participaron 7 candidatos, los seis primeros forman parte del ejercicio que se presenta en este artículo:

1. Alejandra Barrales Magdaleno, de la coalición “Por la CDMX al Frente”, @Ale_BarralesM.
2. Claudia Sheinbaum Pardo, de la coalición “Juntos Haremos Historia”, @Claudiashein.
3. Mikel Arriola, del Partido Institucional Revolucionario (PRI), @MikelArriolaP.
4. Mariana Boy Tamborrell, candidata del Partido Verde Ecologista de México, @mariana_boy.
5. Purificación Carpinteyro Calderón, del Partido Nueva Alianza, @PuriCarpinteyro.
6. Marco Antonio Ignacio Rascón Córdova, del Partido Humanista de la Ciudad de México, @MarcoRascon.
7. Lorena Osornio Elizondo, aspirante sin partido, @LoreOsornio_.

El objetivo principal del ejercicio es mostrar algunos grafos y tablas que son posibles generar con solo algunos datos recuperados de cada perfil. En este caso se presentan 4 diferentes ejemplos en los que se comparan los perfiles de los candidatos con el fin de analizarlos y poder sacar algunas conclusiones acerca de la información que comparten

en sus redes: 1) frecuencia de las palabras utilizadas; 2) frecuencia de las palabras y su proximidad con respecto a otro perfil; 3) palabras distintas más frecuentes entre dos perfiles y 4) palabras con mayor frecuencia en los retuits hechos por los seguidores para un perfil en particular.

Antes de comenzar, es necesario considerar aspectos técnicos relacionados con la extracción de estos a partir de Twitter, así como conocer los principales métodos de procesamiento. A su vez, es importante tener en mente las limitaciones inherentes a este objeto de estudio antes de iniciar una investigación basada en datos de trazo digital (los datos que dejan los usuarios a su paso por las redes sociodigitales). Una restricción para la extracción tiene relación con los API (*Application Programming Interface*) de Twitter, el cual cuenta con dos de ellas para obtener los datos: *streaming* y *REST*. La primera API permite descargar una muestra aleatoria del flujo total de tuits en tiempo real o filtrar por características específicas, y únicamente se tiene acceso al 1% del flujo total de tuits generados. Por su parte, la API REST accede a las bases de datos de Twitter y ofrece información almacenada en ellas. La solicitud a la API REST está limitada a periodos de quince minutos y a cierta cantidad de ítems según el dato solicitado. También, es posible obtener los datos a partir de diferentes herramientas tales como TAGS (<https://tags.hawksey.info/>), The Archivist (<https://es.tweetarchivist.com/>), Gnip (<https://gnip.com/>), NodeXL (<http://nodexl.codeplex.com>), o la librería *twitteR* (<https://cran.r-project.org/web/packages/twitteR/twitteR.pdf>) de R, las cuales combinan extracción, procesamiento y visualización, aunque están sometidas a las restricciones de las API.

Para el presente caso de uso se utilizó la librería *twitteR* de R, lenguaje de programación en el que se pueden aplicar técnicas estadísticas, para realizar la conexión con Twitter y especificar la cuenta de los seis usuarios con el fin de recuperar sus respectivos tuits. Así, el primer paso del proceso consistió en realizar la conexión con Twitter para luego recuperar los tuits a partir de cada perfil. Para el presente ejemplo se utilizan únicamente los tuits generados durante el debate del 18 de abril de 2018.

Para poder trabajar con los datos recuperados de cada perfil es importante que estos estén ordenados (en inglés se nombra a los datos ordenados como *tidy data*), para ello se utilizó la librería de R que se llama *tidytext* [11] y cuyo principal propósito es limpiar los datos antes de proceder a utilizarlos. Usando la librería es posible contar con conjuntos de datos limpios que son fáciles de manipular, modelar y visualizar, y tienen una estructura específica: cada variable es una columna, cada observación es una fila y cada tipo de unidad de observación es una tabla. Los conjuntos de datos ordenados proveen una forma estandarizada para vincular la estructura de un conjunto de datos (diseño físico) con su semántica (su significado) [12].

Los datos ordenados son una forma estándar de mapear el significado de un conjunto de datos a su estructura. Un conjunto de datos puede ser desordenado u ordenado dependiendo de cómo las filas, columnas y tablas coinciden con las observaciones, variables y tipos [12]. En este caso los datos ordenados, que se han extraído de Twitter, responden a:

1. Cada variable forma una columna. Por ejemplo, como variable se tiene el tuit, la fecha de emisión, la localización, el idioma, etc.
2. Cada observación forma una fila. En este sentido, una observación corresponde a cada tuit realizado, así como todos sus atributos o variables.

3. Cada tipo de unidad de observación forma una tabla. En este caso corresponde a cada uno de los perfiles de los candidatos.

El conjunto de datos es una colección de valores, usualmente pueden ser tanto números o cadenas de caracteres. Cada valor pertenece a una variable y a una observación. Por su parte, una variable, contiene todos los valores que miden al mismo atributo subyacente en unidades (número de seguidores, número de respuestas, etc.). Mientras tanto, una observación contiene todos los valores medidos en la misma unidad (como el nombre del perfil, a quién está dirigido el mensaje, o la fecha) a través de los atributos.

En la Fig. 1, se muestra una tabla para el perfil, donde la unidad de observación corresponde a @Ale_BarralesM. En este caso se han extraído 19 parámetros los cuales corresponden a las variables: *text*, *favorited*, *favoriteCount*, *replyToSN*, *created*, *truncated*, *replyToSID*, *id*, *replyToUID*, *statusSource*, *screenName*, *retweetCount*, *isRetweet*, *retweeted*, *longitude*, *latitude*, *location*, *language*, y *profileImageUrl*. Se extrae una tabla, unidad de observación, para cada uno de los perfiles.

text	favorited	favoriteCount	replyToSN	created
Es inaceptable que hoy, @Claudiashein nos hable com...	FALSE	217	Ale_BarralesM	2018-04-19 01:47:11
Mientras la ciudad necesitaba agua, @Claudiashein la ...	FALSE	113	Ale_BarralesM	2018-04-19 01:40:32
.@Claudiashein es la <U+00FA>nica Secretaria de Me...	FALSE	147	Ale_BarralesM	2018-04-19 01:39:01
Estoy obligada a se<U+00F1>alar que @Claudiashein...	FALSE	330	NA	2018-04-19 01:31:28

Fig. 1. Ejemplo de tabla, unidad de observación, para @Ale_BarralesM antes de aplicar *tidytext* para ordenar los datos.

Posteriormente, se aplica la librería *tidytext*, ver Fig. 2, para trabajar con datos ordenados y solo con aquellos que se necesitarán para el análisis posterior. En este caso, los atributos seleccionados y que forman parte de los datos ordenados son: *screenName*, *created*, *retweetCount* y *favoriteCount*. En la Fig. 2 se muestra un ejemplo con el perfil de @Ale_BarralesM aunque la tabla creada contiene a todos los candidatos.

text	screenName	created	retweetCount	favoriteCount
#PorLaCDMXalFrente seguiremos luchando por la equ...	Ale_BarralesM	2018-01-10 21:19:39	14	30
#PorLaCDMXalFrente seguiremos trabajando para que...	Ale_BarralesM	2018-01-22 01:16:46	20	56
#PorLaCDMXalFrente vamos a crear las condiciones p...	Ale_BarralesM	2017-12-29 01:30:07	24	68
#PorLaCDMXalFrente vamos a fortalecer los programa...	Ale_BarralesM	2018-02-03 02:28:52	45	106

Fig. 2. Ejemplo de tabla con *tidytext* donde todos los perfiles están ordenados con las variables.

Para el análisis es importante precisar que no es de interés analizar los tuits si están repetidos, es decir que si son retuits enviados por los seguidores con una vez que se analicen es suficiente. Por lo tanto, uno de los primeros pasos es eliminar los retuits a su vez que se quitan las palabras vacías. Sin embargo, es importante conservar los *hashtags* ya que permitirán realizar un análisis posterior.

Uno de los clásicos análisis que se realizan con texto corresponde a calcular la frecuencia de palabras, es decir cuántas veces cada uno de los candidatos uso cada palabra que aparece en sus tuits. En este caso se presenta, en la Fig. 3, el número total de palabras diferentes utilizadas por cada uno de los actores.

screenName	word	n	total	freq
Ale_BarralesM	frente	154	41167	0.003740860
Ale_BarralesM	buen	152	41167	0.003692278
Ale_BarralesM	mujeres	152	41167	0.003692278
MarcoRascon	como	150	30809	0.004868707
Claudiashein	colonia	149	29545	0.005043155
Ale_BarralesM	buenos	148	41167	0.003595113
MarcoRascon	hoy	147	30809	0.004771333
MarcoRascon	independiente	146	30809	0.004738875

Fig. 3. Ejemplo de la tabla que muestra el total y el porcentaje de frecuencia para cada una de las palabras utilizadas por cada candidato.

Una de las razones de utilizar R está en su gran potencial para generar gráficos que permiten visualizar los resultados obtenidos. Por ejemplo, en la Fig. 4, se hace una comparación del número de distintas palabras empleadas entre los candidatos Claudia Sheinbaum (@Claudiashein) y Mikel Arriola (@MikelArriolaP).

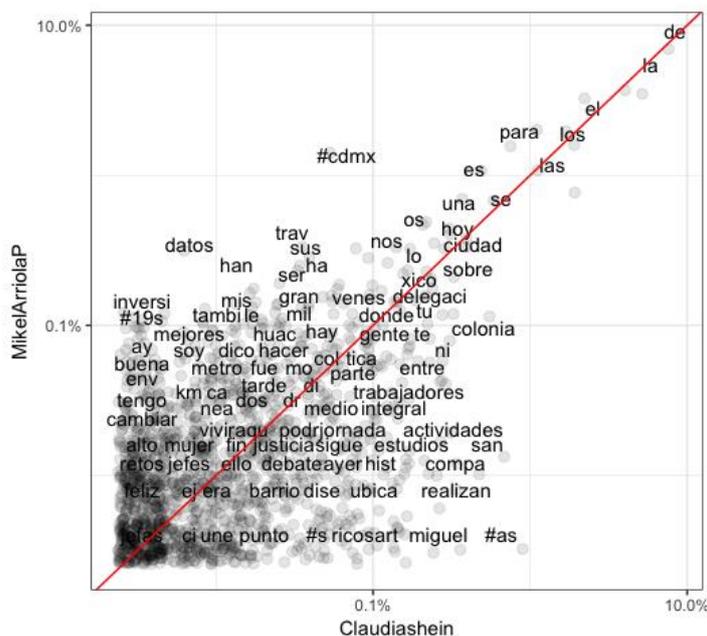


Fig. 4. Comparación en el uso de palabras entre dos de los candidatos.

Las palabras que están cercanas a la línea, en la Fig. 4, tienen frecuencias similares en ambos corpus de los candidatos. Por ejemplo, “fin”, “aquí”, “gente”, “hoy”, “ciudad”. En cambio, las palabras que están lejos de la línea se encuentran más en un

perfil que en el otro. Por ejemplo, Mikel Arriola aborda más el caso de “#19s” o “#cdmx” que Claudia Sheinbaum.

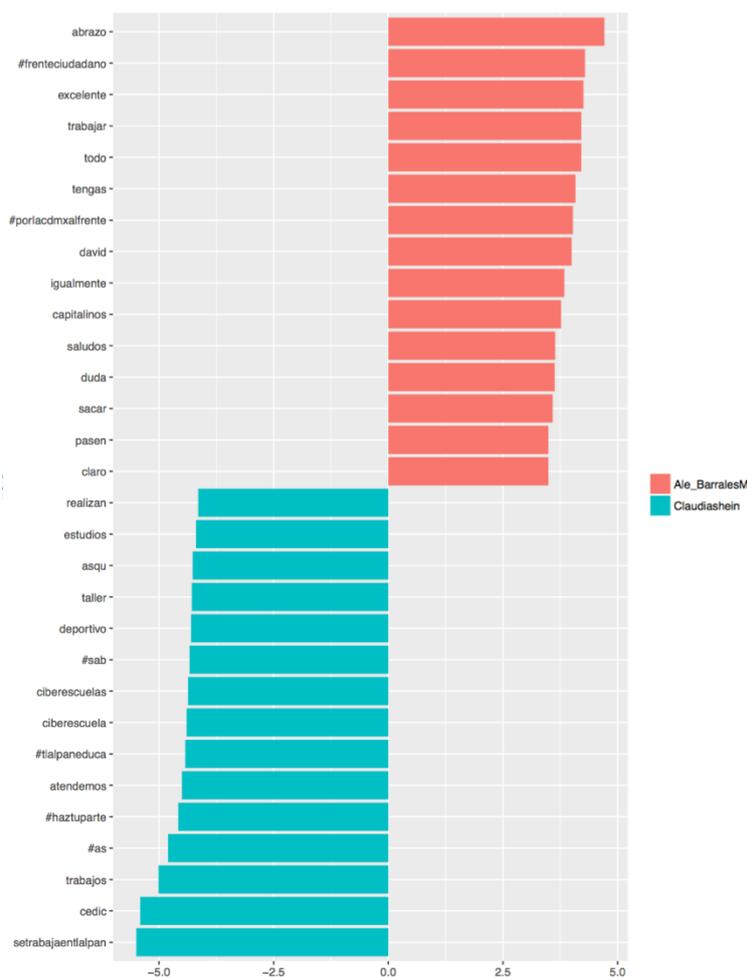


Fig. 5. Comparación en el uso de palabras entre las candidatas Alejandra Barrales y Claudia Sheinbaum.

Para poder obtener mayor información acerca del contenido que cada candidato publica, es posible saber cuáles son las 15 palabras más distintivas para cada cuenta. En las Fig. 5 y 6 se muestran dos diferentes ejercicios de comparación. En el primero, se utilizaron las cuentas de @Ale_BarralesM y @Claudiashein en el que se aprecia por una parte las palabras de “#frenteciudadano” y “#porlacdmxalfrente” característicos de la candidata Alejandra Barrales. En el caso de la candidata Claudia Sheinbaum destaca su trabajo en la Delegación Tlalpan con las palabras “#tlalpaneduca”, “setrabajaentlalpan” o “cedic” que se refiere a uno de los centros comunitarios.

Con el análisis de la Fig. 6, es posible vislumbrar diferentes problemas abordados, respectivamente, por cada uno de los perfiles. En el caso del candidato Mikel Arriola, al igual que Claudia Sheinbaum, están presentes los temas de la delegación, así como el agua, el sismo, y Xochimilco. Por su parte, las palabras características del discurso de Marco Rascón están representadas con: “#humanicemoslaciudad”, “constituyente”, “independiente”, “asamblea”, “cnte”, “cultura” y hasta “Trump”.

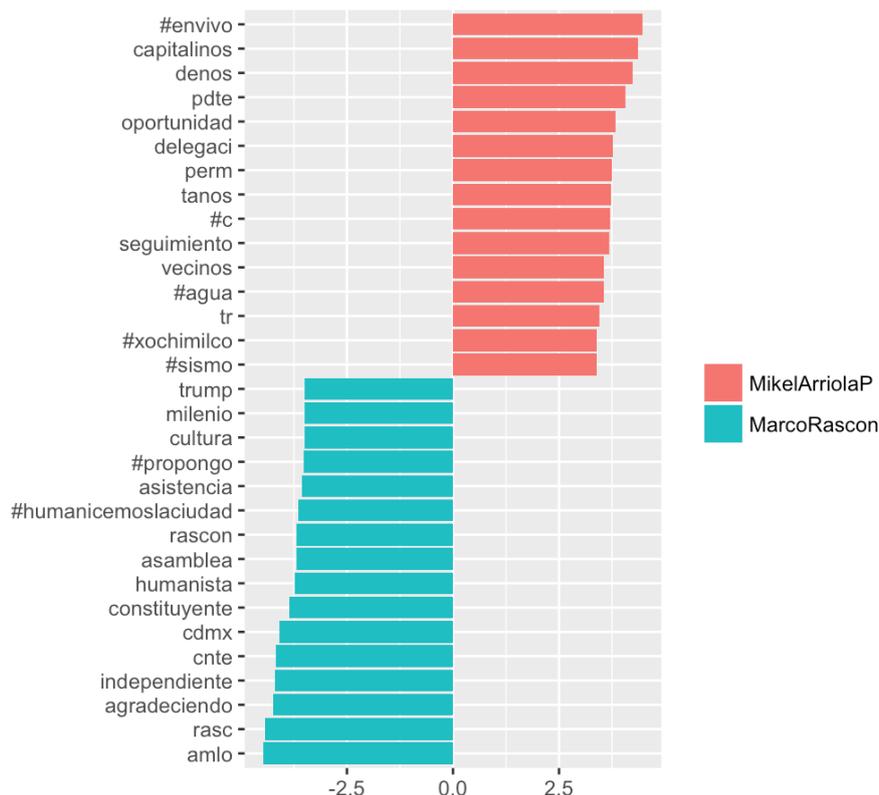


Fig. 6. Comparación en el uso de palabras entre los candidatos Mikel Arriola y Marco Rascón.

Por lo regular, el usuario que sigue, lee y apoya a los candidatos es proclive a compartir, en forma de retuit, información que considera importante e interesante. Sin embargo, hay algunos tuits que son de mayor interés para los usuarios y que pueden ser detectados a partir de las palabras que usan. En el siguiente ejemplo, se muestran las palabras que están contenidas en una mayor cantidad en los tuits compartidos, retuiteados. Para ello, para cada perfil se extrajo el número total de tuits que fueron a su vez retuiteados. Por ejemplo, en la Fig. 7, las palabras que más se encuentran en los retuits de la candidata Purificación Carpinteyro (@PuriCarpinteyro) son: “encuentro”, “neta”, “candidatos”, “trabajo”, “apoyo”, “momento”, “nuestra”, “campa”, “vamos” y “propuestas”. La gráfica muestra la mediana de retuits efectuados a partir de la cual se puede conocer, también, la campaña hecha por los propios ciudadanos en apoyo a su respectivo candidato.

A partir de los ejemplos previos es importante destacar que no se muestra la comparativa de todos los perfiles pero que, en algunos casos, como en la Fig. 7, se puede concluir acerca de quiénes tienen un mayor número de seguidores y apoyo en las redes. Desde luego, para una conclusión así es deseable un análisis más profundo sobre quiénes son estos seguidores y en qué casos los *bots* hacen de las suyas.

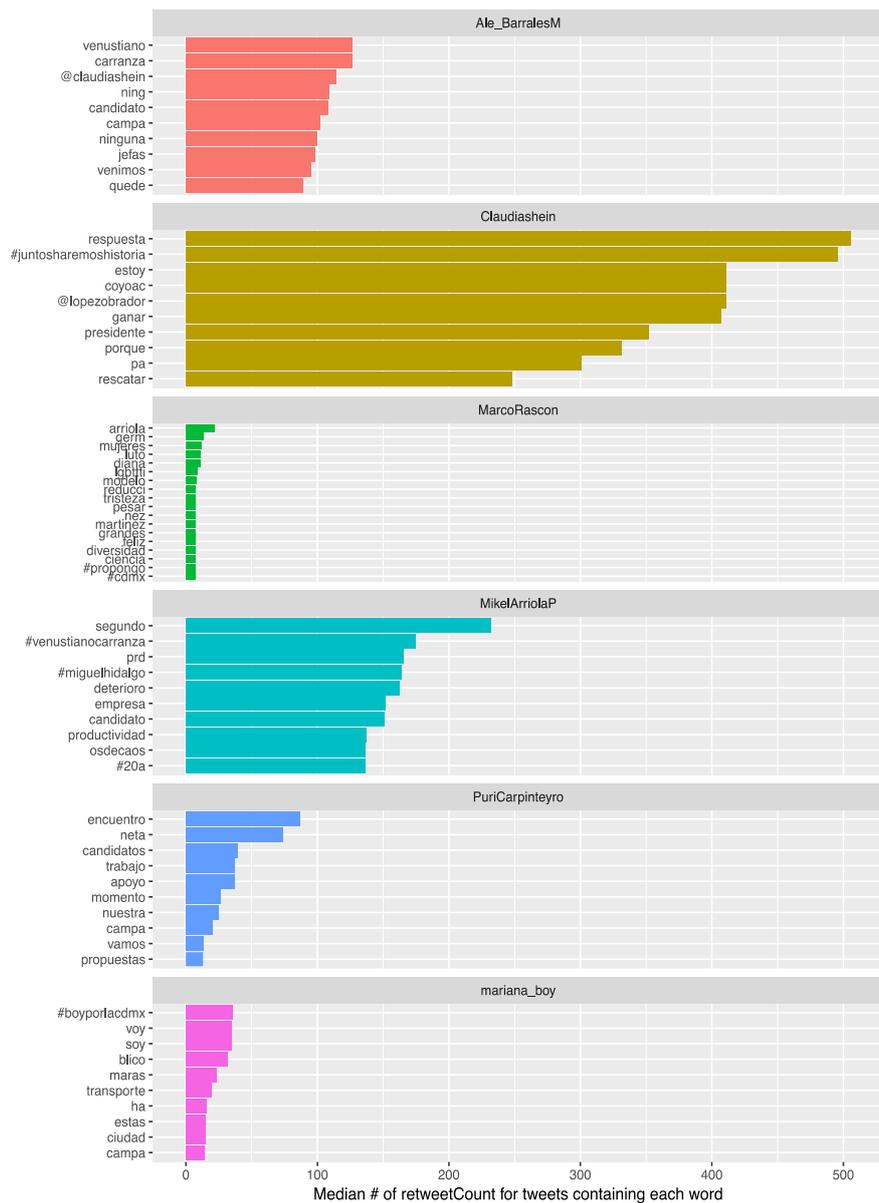


Fig. 7. Palabras que aparecen con mayor frecuencia en los retuits que hacen los ciudadanos de cada uno de los candidatos.

4. Discusión

Al trabajar con datos que provienen de redes sociodigitales es importante considerar que estos se generan a un ritmo considerable y que su manipulación requiere de técnicas que permitan pasar de un formato semiestructurado a uno estructurado.

El uso de principios de ordenamiento de datos puede hacer que muchas tareas de minería de textos sean más fáciles, más efectivas y consistentes con las herramientas que ya están en uso. El ordenamiento de los datos permite manipular, resumir y visualizar las características del texto fácilmente e integrar el procesamiento del lenguaje natural en flujos de trabajo efectivos.

Los datos ordenados permiten que su manipulación sea más fácil al trabajar sabiendo que toda la información con la se cuenta tiene la misma estructura y cuenta con las mismas características. Así es posible realizar diferentes acciones tales como: 1) filtrar el contenido para trabajar solo con ciertas variables; 2) unir o colapsar múltiples valores; y, por ejemplo, 3) ordenar para permitir una visualización de la información pertinente.

A partir de los ejercicios anteriores es posible apreciar algunos de los posibles análisis que pueden realizarse a partir de la recuperación de tuits de determinados perfiles. En estos ejercicios se ha realizado lo siguiente:

1. Para cada perfil se contabilizó la frecuencia de las palabras utilizadas (ver Fig. 3).
2. Para dos perfiles seleccionados se comparó la frecuencia de las palabras y su proximidad (ver Fig. 4).
3. Para dos perfiles distintos, se compararon las palabras frecuentes más distintas entre ellos (ver Fig. 5 y 6).
4. Para cada perfil, se encontraron las palabras que aparecen con mayor frecuencia en los retuits hechos por sus seguidores (ver Fig. 7).

Se ha utilizado únicamente el aspecto sintáctico de las palabras que aparecen en los tuits debido a que se considera que la gran mayoría de estos provienen del mismo contexto por lo que no es necesario explorar su semántica. Sin embargo, algunas de las limitaciones a las que los investigadores se enfrentan al trabajar con tuits tienen que ver principalmente con 1) la inmediatez de los tuits generados; 2) la gran cantidad de tuits, que al menos en contextos políticos, se generan; 3) la ironía de los tuits, la cual sigue siendo un gran reto a la hora de identificar el sentido del tuit y 4) el contexto en el que se generan los tuits que pueden proveer mayor información sobre qué está diciendo cada candidato.

5. Conclusiones

En este artículo se presenta una breve introducción al tipo de análisis que puede realizarse con datos provenientes de Twitter. Debido a la gran velocidad con la que circula la información en las redes sociodigitales, es importante contar con técnicas que permitan que el trabajo de recuperación, extracción, análisis y visualización sea más rápido y sencillo. Para ello, se utiliza R como lenguaje de programación al ofrecer librerías que permiten agilizar la minería de textos, tal como lo hace *tidytext*.

La interacción entre los políticos y los ciudadanos adquiere muchas formas, dependiendo del contexto y espacio donde tenga lugar. En el ambiente digital, ocurre principalmente en medios sociodigitales. Actualmente, Twitter permite recuperar los tuits para poderlos analizar y poder observar a detalle distintas configuraciones comunicativas para estudiarlas, considerando su principal activo en los usuarios y sus redes.

Una comprensión más profunda de la actividad en Twitter se lograría a través del análisis del contenido o del sentimiento expresados en los tuits. Incluir esta dimensión semántica ayudaría en la depuración de los datos extraídos. El análisis puede ser complementado con la sistematización de acciones, como la búsqueda de patrones estructurales a través de distintas redes y la detección de micro-eventos –sucesos relevantes dentro de cada evento– mediante la comparación de fechas de publicación. Asimismo, el análisis del contenido difundido mediante Twitter si bien se hace mediante algoritmos computacionales debería ser complementado con un equipo interdisciplinario de manera a obtener un panorama más completo sobre la implicación, influencia y determinación que tiene la comunicación política a través de redes sociodigitales.

Referencias

1. Christakis, N.A. (et. al.): Conectados: el sorprendente poder de las redes sociales y cómo nos afectan. Pensamiento (Taurus (Firm)). Aguilar, Altea, Taurus, Alfaguara, Santillana Ediciones Generales, S.A. de C.V., 368 p. (2010)
2. Yusuf, H.: Old and New Media: Converging During the Pakistan Emergency (Marzo 2007–Febrero 2008), 61 p., Cambridge, Mass.: Massachusetts Institute of Technology, Center for Future Civic Media. <http://civic.mit.edu/blog/humayusuf/old-and-new-media-converging-during-the-pakistan-emergency-march-2007-february-2008>. Último acceso: 07/05/2019
3. Zimmer, M., Proferes, N.: A topology of Twitter: disciplines, methods, and ethics. *Aslib Journal of Information Management* 66(3), 250–261. DOI: 10.1108/AJIM-09-2013-0083 (2014)
4. Vyas, V., Uma, V.: Open Issues in Opinion Mining. *Extracting Knowledge from Opinion Mining*, p. 283. (2018)
5. Fan, R. (et. al.): Anger is more influential than joy: sentiment correlation in Weibo. *PLoS One* 9(10) e110184. DOI: 10.1371/journal.pone.0110184 (2013)
6. Castells, M.: *Redes de Indignación y Esperanza- Los movimientos Sociales en la Era de Internet*. (1ra. Ed) Madrid: Alianza Editorial, S. A (2012)
7. Cruz, A.: Esfera pública digital: viejas prácticas en 140 caracteres. *Twitter en las elecciones México*. *VIRTUalis* 5(9), 76–89 (2015)
8. Cano, C.: Las redes, las calles y los medios. Análisis visual de las protestas del #1Dmx 2014 en Twitter-México. *VIRTUalis* 6(11), 49–72 (2015)
9. Reyes, J.M.: Twitter, tuits y tuiteros: Análisis de una sentencia en materia electoral. *Juez. Cuadernos de Investigación* (5) (2016)
10. Suárez-Serrato, P. (et al.): On the influence of social bots in online protests. In: *International Conference on Social Informatics*. Springer, Cham, 2016, pp. 269–278 (2016)
11. Silge, J., Robinson, D.: tidytext: Text Mining and Analysis Using Tidy Data Principles in R. *JOSS* 1 (3). The Open Journal. <https://doi.org/10.21105/joss.00037> (2016)
12. Wickham, H., (et al.): Tidy data. *Journal of Statistical Software* 59(10), 1–23 (2014)