

“Influencia de medios digitales en la opinión pública reflejada en twitter”

Kevin Daniel Molina Bejar
Ciencias de la computación, UNI
Lima, Perú
kevin.molina@uni.pe

Abstract— El propósito de este estudio es encontrar el alcance que tienen los medios digitales en la opinión de la gente, ya sea provocando opiniones a favor o en contra del contenido difundido, usando la plataforma de twitter como dato en tiempo real.

I. INTRODUCCIÓN

Los medios digitales juegan un rol importante en el desarrollo de nuestra sociedad dado que pasamos la mayor parte del tiempo siempre conectados a las noticias o tendencias de la actualidad, ellos eligen qué contenidos darles más difusión o no y esto ejerce una gran influencia en la opinión pública la que en nuestro país tiene 22 millones (68%) de usuarios de internet, sobre los 32.3 millones de la población total. Los usuarios de redes sociales también alcanzan los 22 millones.¹

En el Perú existe un gran grupo mediático (“grupo el comercio”) que domina fuertemente la circulación y lectoría mediante sus 9 diarios de alcance local y nacional, también el tráfico web a través de sus 13 medios digitales de noticias más importantes.²

Twitter es una de las plataformas más populares que te permite consultar tendencias, buscar, publicar, compartir información o actualizaciones en muy pocos pasos. Por eso elegimos twitter como una buena medida en tiempo real de las opiniones o sentimientos de las personas.

El impacto de los contenidos con mayor difusión en medios digitales se puede ver reflejado en las opiniones que se emiten vía twitter ya sean positivas o negativas así como la rápida difusión de algún viral o noticia. Esto es aprovechado por los bloques políticos o candidatos a puestos gubernamentales para hacer tratos bajo la mesa y así manipular los contenidos que generan debates en nuestra sociedad.

Por estas razones es que buscaremos la relación entre los sentimientos de los tuits de las personas con los contenidos del grupo que tiene la mayor concertación medios digitales.

II. ESTADO DEL ARTE

Se encontraron los siguientes paper relacionados con el tema para partir de ahí nuestra propuesta:

A. Predicting the Future With Social Media[1]

1) *Autores:* Sitaram Asur and Bernardo A. Huberman, Social Computing Lab.

2) *Objetivo general:* Este paper muestra como un modelo simple construido a partir de la velocidad a la que se crean los tweets sobre temas particulares puede superar a los predictores basados en el mercado. Además, demuestra cómo los sentimientos extraídos de Twitter se pueden utilizar aún más para mejorar el poder de previsión de las redes sociales.

3) *Resultados:* En este paper se enfocaron en el problema de predecir los ingresos de taquilla de las películas por el hecho de tener una métrica clara de comparación con otros métodos, este método puede extenderse a una gran variedad de temas, que van desde la calificación futura de los nuevos productos hasta los resultados de las elecciones.

B. Twitter y la teoría de la Agenda-Setting: mensajes de la opinión pública digital [2]

1) *Autor:* Raquel Rubio García, Universidad Carlos III de Madrid.

2) *Objetivo general:* Este paper pretende comprobar si la teoría de la AgendaSetting mantiene su vigencia dentro del creciente entorno digital; concretamente dentro del contexto de las redes sociales, para este caso Twitter en España. Se estudia cómo los medios de comunicación tradicionales marcan los temas más comentados por los usuarios de Twitter España.

3) *Resultados:* En este paper se desprende que existe una clara correspondencia entre la agenda del público y la establecida por los medios de comunicación tradicionales. Dejando al margen la parte de la agenda pública referida a los temas originarios de Twitter, se observa que en lo que respecta a las temáticas más informativas, estas han encontrado su reflejo en los medios analizados, es decir, que los asuntos comentados por los usuarios de Twitter en la red social han sido en su mayoría abordados también por los medios.

C. Detecting Campaign Promoters on Twitter using Markov Random Fields [3]

1) *Autores:* Huayi Li, Arjun Mukherjee, Bing Liu, Rachel Kornfield and Sherry Emery, Department of Computer Science Institute for Health Research and Policy University of Illinois at Chicago, IL, USA. Department of Computer Science University of Houston, TX, USA.

2) *Objetivo general:* Las campañas en las redes sociales a menudo aparecen en modos sigilosos. Los promotores de

¹ (n.d.). Digital Report 2018 - We Are Social. Se recuperó el septiembre 13, 2018 de <https://digitalreport.wearesocial.com/>

² (n.d.). Concentración | Media Ownership Monitor - Peru. Se recuperó el septiembre 13, 2018 de <https://peru.mom-rsf.org/es/hallazgos/concentracion/>

campañas a menudo intentan influir en las conductas / opiniones / decisiones de las personas de manera latente, de modo que los lectores no son conscientes de que los mensajes que ven son publicaciones estratégicas para persuadirlos a comprar productos / servicios específicos. Dado un conjunto de tweets transmitidos desde Twitter basados en un conjunto de palabras clave que representan un tema en particular, la técnica propuesta tiene como objetivo identificar las cuentas de los usuarios que están involucradas en la promoción de estas campañas ocultas en twitter.

3) *Resultados*: Usando Markov Random Fields se alcanza una precisión de 0.91.

III. DISEÑO DEL EXPERIMENTO

El experimento contará con 4 partes:

- Recolectar todo el texto de las noticias de los medios digitales.
- Clasificar los temas de las noticias con mayor frecuencia publicadas por los medios de comunicación.
- Recolectar todos los tuits relacionados con lo obtenido en el paso anterior.
- Clasificaremos lo recolectado según los sentimientos en positivo o negativo.

A. *Scraping web*

El scraping web o raspado, consiste en extraer toda la data que necesites de alguna página web y darle forma de modo que se entienda y de acuerdo a tus necesidades, para nuestro caso haremos uso de las librerías Requests y BeautifulSoup para obtener todas las noticias publicadas por los medios digitales en sus respectivas páginas web y guardarlas en una base de datos alojada en una instancia de AWS.

B. *Natural language processing*

Se hará uso de 2 librerías de python para la clasificación de texto, scikit learn³ y NLTK(Natural Language Toolkit)⁴.

C. *API twitter*

Se usará la api de código abierto de twitter para recolectar los tuits de las personas respecto a cada tema consultado y así poder clasificarlos como comentarios positivos o negativos con ayuda del natural language processing.

³ (n.d.). Working With Text Data — scikit-learn 0.19.2 documentation. Se recuperó el septiembre 14, 2018 de http://scikit-learn.org/stable/tutorial/text_analytics/working_with_text_data.html

⁴ (n.d.). NLTK Book. Se recuperó el septiembre 14, 2018 de <https://www.nltk.org/book/>

REFERENCIAS

- [1] [Predicting the Future With Social Media, 2010.](#)
- [2] [Twitter y la teoría de la AgendaSetting: mensajes de la opinión pública digital, 2013](#)
- [3] [Detecting Campaign Promoters on Twitter using Markov Random Fields 2014](#)