

# ElectionMap: una representación geolocalizada de intenciones de voto hacia partidos políticos sobre la base de comentarios de usuarios de Twitter

*ElectionMap: a geolocalized representation of voting intentions to political parties based on twitter's user comments*

Francisco Agulló, Antonio Guillén, Yoan Gutiérrez, Patricio Martínez-Barco  
Carretera San Vicente del Raspeig S/N, 03690 San Vicente del Raspeig  
Departamento de Lenguajes y Sistemas Informáticos, Universidad de Alicante  
{fagullo,aguillen,ygutierrez,patricio}@dlsi.ua.es

**Resumen:** ElectionMap es una aplicación web que realiza un seguimiento a los comentarios publicados en *Twitter* en relación a entidades que refieren a partidos políticos. Las opiniones de los usuarios sobre estas entidades son clasificadas según su valoración y posteriormente representadas en un mapa geográfico para conocer la aceptación social sobre agrupaciones políticas en las distintas regiones de la geografía española.

**Palabras clave:** Twitter, análisis de sentimientos, representación en mapas

**Abstract:** ElectionMap is a web application that follows, in *Twitter*, entities previously established and related to the politics. The user's opinions about the entities are classified according to its valuation by using sentiment analysis processes. Afterwards the opinions are represented in a geographic map that allows to know the social acceptance of spanish political parties in different geographical areas.

**Keywords:** Twitter, sentiment analysis, map representation

## 1 Introducción

En la actualidad las redes sociales se han convertido en uno de los principales medios de distribución de información. La cantidad de información que se genera es tan grande que tanto empresas como gobiernos han comenzado a guiar sus campañas de promoción fijándose en las opiniones que los usuarios de las redes sociales valoran en sus perfiles. Una de las redes más utilizadas en la actualidad es *Twitter*, y la cantidad de información que proporciona esta red social no ha pasado desapercibida para las grandes empresas y gobiernos.

Otro de los problemas de las campañas de promoción es que no tienen el mismo impacto en toda la geografía. Saber en qué regiones una campaña obtiene una mayor aceptación social.

Para cubrir estas necesidades surge *ElectionMap*. Esta herramienta muestra de forma gráfica la opinión de los usuarios de *Twitter* sobre temas relacionados con la política. Como se describe en la sección 3, no todas las opiniones pueden tenerse en cuenta, sólo sir-

ven las opiniones positivas y localizadas, por lo que el conjunto de datos utilizado no es el 100 % de la información recogida, si no que una muestra representativa del total. Finalmente tras procesar este conjunto de datos se muestra una representación gráfica de intención de voto político agrupada por áreas geográficas representando así el apoyo de los usuarios de los medios sociales. La aplicación web se divide en dos componentes, el núcleo que se explica en la sección 2 y la parte web que se detalla en la sección 3. Por último en la sección 4 se puede ver el procedimiento utilizado para la detección de opiniones.

## 2 Núcleo

El núcleo es el encargado de recopilar la información de los tweets en los cuales se mencionan los partidos a evaluar. Para ello, se define un conjunto de términos para cada una de las entidades que se van a representar. Cada vez que un usuario de *Twitter* escribe uno de los términos que identifican a los partidos a evaluar, se lanza una alerta hacia nuestro sistema notificando el nuevo comentario y se almacena la información relevante al mensa-

je. Además de la información que *Twitter* nos proporciona, para cada uno de los tweets, se realiza un procesamiento del tweet para conocer si el texto del mensaje es positivo, negativo o neutral y se almacena para considerarlo al igual que el resto de la información.

Por otro lado, el núcleo también se encarga de filtrar la información obtenida de *Twitter* ya que se reciben muchos datos pero no todos son útiles para mostrarse geoespacialmente ya sea porque el usuario no tiene activada la geolocalización o porque en su perfil no detalla su origen. Además al ser un mapa para valorar la intención de voto existen una serie de restricciones que vienen impuestas por el objetivo:

- **Los votantes sólo suman, no restan.** Los usuarios pueden expresar tanto su conformidad como su disconformidad con los partidos políticos, ya que cuando un votante acude las urnas vota en positivo, no en negativo, todos los mensajes con valoraciones negativas son ignorados y no quedan representados en el mapa.
- **Un usuario un voto.** Un usuario puede opinar bien sobre más de un partido político, pero en unas elecciones sólo podría formalizar el voto sobre una de las formaciones políticas, por lo que se asume que un usuario votará a la formación que mayor valoración haya obtenido en el total de sus mensajes publicados en la red social, lo cual indicaría que este usuario tiene mayor inclinación sobre ese partido al que ha comentado positivamente en más ocasiones.
- **Localización del tweet.** Los datos se muestran en un mapa de España, por lo que es necesario saber a que localidad le corresponde el voto de cada usuario. Para ello en primer lugar se busca en la descripción de usuario, si ésta tiene una localización válida, se asume que esa es la localización real del usuario. En caso de no obtener ninguna localización en la búsqueda, se comprueba si el usuario tiene activada la geolocalización. En caso de tener algún tweet con localización almacenado, se tiene en cuenta la localización que más veces se ha utilizado, y si no hay ningún tweet geolocalizado, se ignoran los mensajes del usuario ya que no se podrían situar en el mapa.

- **Formaciones con restricciones.** Por último algunas de las formaciones políticas sólo se presentan en ciertas regiones, pero en *Twitter* puede haber usuarios que valore a estos partidos desde regiones en las que no sería posible su votación. Para evitar contabilizar estas opiniones, se permite añadir filtros en regiones, de forma que una formación puede estar filtrada a una o más regiones. Además un filtro por región se hereda a las divisiones territoriales de nivel inferior, es decir, un filtro de comunidad se extenderá a todas las provincias de la comunidad y éste a su vez a todas las ciudades de las provincias.

### 3 Aplicación web

La aplicación web es la parte visual de la herramienta la cual se encuentra disponible públicamente en la siguiente dirección web: <http://gplsi.dlsi.ua.es/demos/electionmap>. Para la representación de los datos se ha utilizado el API de Google Maps<sup>1</sup> al que se han añadido una serie de controles propios. En la figura 1 podemos ver las tres secciones de la aplicación que a continuación se van a detallar.

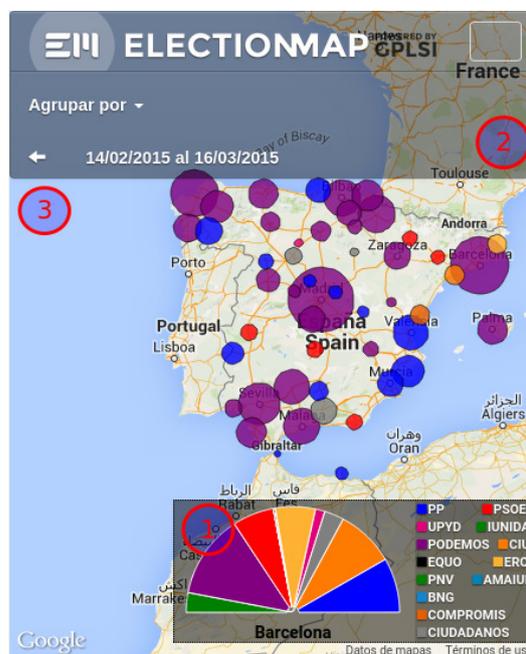


Figura 1: Vista de la aplicación.

- **Sección 1.** En este control se puede observar un gráfico (ordenado igual que

<sup>1</sup><https://developers.google.com/maps/?hl=es>

el parlamento) con la intención de voto para cada partido en una determinada área geográfica. El gráfico se actualiza automáticamente cuando se selecciona alguna de las burbujas de la *Sección 3*.

- **Sección 2.** Menú de la aplicación, en el podemos editar tanto el modo de visualización como el periodo de tiempo que queremos ver representado en el mapa. Ya que se trata de un mapa español, podemos distinguir 4 modos para visualizar los datos:
  - Agrupados por ciudades.
  - Agrupados por provincias.
  - Agrupados por comunidades.
  - Agrupados por país.
- **Sección 3.** El último de los controles es en el que se muestran las burbujas. Una burbuja representa el color de la formación política que más valoraciones positivas tiene en una área geográfica. Además el tamaño de cada una de las burbujas es proporcional al porcentaje de usuarios con intención de voto hacia esa formación política obtenidos así como al número de habitantes del área representada. Por último si se quiere dejar una burbuja fijada en el gráfico de la *Sección 1* basta con hacer *click* sobre dicha burbuja y la información quedará fijada por un intervalo de 10 segundos. Si se desea desfijar antes de que este intervalo finalice se puede hacer *click* en cualquier zona del mapa.

#### 4 *Análisis de opiniones*

Como se ha comentado anteriormente, la aplicación realiza un proceso de recogida y de almacenamiento de datos, pero entre estos dos procesos se realiza un procesamiento intermedio en el cual se evalúa si un texto expresa apoyo o rechazo hacia las entidades (ii.e. partidos políticos) a valorar. El método de evaluación utilizado es el comentado en “*The OpAL System at NTCIR 8 MOAT*” (Balahur et al., 2010a). Este método ya se demostró con éxito cuando se aplicó a otras tareas de la minería de opiniones como en el caso del Opinion Question Answering (Balahur et al., 2010b). En los artículos anteriormente citados se puede encontrar una explicación

detallada sobre los léxicos que se combinan así como el uso de modificadores léxicos y elementos de potenciación de características utilizados.

El método se basa en un conjunto de lexicones que contienen palabras que expresan opiniones positivas o negativas agrupados según su intensidad. Cuando se evalúa un texto, se buscan las palabras contenidas en los lexicones y se les asigna una valoración base según el lexicon en el que se encuentren. El método además utiliza *stemmer* para valorar familias de palabras que pudieran estar relacionadas con las distintas palabras de los lexicones. Las palabras que no aparecen en ninguno de los lexicones se consideran neutrales. Este valor puede ser modificado posteriormente según los potenciadores y los modificadores que se encuentren en la frase. *OpAL* tiene en cuenta distintos tipos de modificadores (ii.e muy, mucho, poco, menos, no, ningún, etc) que inciden en el cambio o modificación de la polaridad de las distintas palabras con carga sentimental implicadas en los mensajes procesados. Finalmente se suman las valoraciones de cada una de las palabras y el valor resultante es el que se utiliza para etiquetar el texto como positivo (valoraciones mayores que 0), neutral (valoraciones iguales a 0) o negativo (valoraciones inferiores a 0).

#### 5 *Agradecimientos*

*ElectionMap* es una aplicación web desarrollada por el Grupo de Procesamiento del Lenguaje Natural y Sistemas de Información (GPLSI)<sup>2</sup> de la Universidad de Alicante<sup>3</sup>. Esta aplicación ha sido parcialmente financiada por el Gobierno Español y la Comisión Europea a través de los proyectos: AT-TOS (TIN2012-38536-C03-03), LEGOLANG (TIN2012-31224), SAM (FP7- 611312) y FIRST (FP7-287607) y por la Universidad de Alicante a través del proyecto emergente “Explotación y tratamiento de la información disponible en Internet para la anotación y generación de textos adaptados al usuario” (GRE13-15).

#### *Bibliografía*

Balahur, A., E. Boldrini, A. Montoyo, y P. Martínez-Barco. 2010a. The opal system at ntcir 8 moat. En *Proceedings of*

<sup>2</sup><http://gplsi.dlsi.ua.es/>

<sup>3</sup><http://www.ua.es/>

*NTCIR-8 Workshop Meeting, Tokyo, Japan*, páginas 241–245.

Balahur, A., E. Boldrini, A. Montoyo, y P. Martínez-Barco. 2010b. Opinion question answering: Towards a unified approach. En *ECAI*, páginas 511–516.