

Técnicas para el análisis del sentimiento en Twitter: Aprendizaje Automático Supervisado y SentiStrength

Tecnhiques for sentiment analysis in Twitter: Supervised Learning and SentiStrength

Tomás Baviera

tobapui@upv.es

Universitat Politècnica de València

<http://orcid.org/0000-0002-2331-6628>

Recibido: 25/10/2016

Aceptado: 05/12/2016

RESUMEN

El análisis del sentimiento en los mensajes publicados en Twitter ofrece posibilidades de gran interés para evaluar las corrientes de opinión difundidas a través de este medio. Los enormes volúmenes de textos requieren de herramientas capaces de procesar automáticamente estos mensajes sin perder fiabilidad. Este artículo describe dos tipos de técnicas para abordar este problema. La primera estrategia se basa en los procesos de Aprendizaje Automático Supervisado. Su aplicación requiere integrar algunas herramientas del Procesamiento de Lenguajes Naturales y tomar como punto de partida un corpus clasificado. El segundo enfoque está basado en diccionarios de polaridad. En esta línea se sitúa la herramienta de SentiStrength, la cual se está aplicando cada vez más a los estudios de Twitter en inglés. El artículo evalúa los estudios más avanzados que utilizan cada uno de estos enfoques para el análisis de los tweets en castellano. Por último, se señalan las ventajas y limitaciones de cada uno de estos enfoques para su aplicación a la investigación en comunicación política. Si bien el aprendizaje automático supervisado permite tener en cuenta el contexto, el investigador requiere competencias de analista de datos con el fin de afinar mejor el proceso. En cambio, SentiStrength está más orientado al contenido semántico de los términos del mensaje, y se requiere más bien una competencia en lingüística por parte del investigador. La principal conclusión es que ambos métodos automáticos de análisis no pueden prescindir de una exigente codificación manual si se desea utilizarlos con fiabilidad en la investigación.

PALABRAS CLAVE

comunicación política, Twitter, análisis de sentimiento, aprendizaje automático supervisado, SentiStrength

ABSTRACT

Sentiment analysis on Twitter offers possibilities of great interest to evaluate the currents of opinion disseminated through this medium. The huge volumes of texts require tools able to automatically process these messages without losing reliability. This paper describes two different types of approaching this problem. The first strategy is based on Supervised Learning processes, developed in the field of artificial intelligence. Its application requires some tools from natural language processing along with a classified corpus as a starting point. The second approach is based on polarity dictionaries. SentiStrength tool is located in this line. It is increasingly applied to studies of Twitter in English. The paper assesses the most advanced studies using each of these approaches for analyzing tweets in Spanish. Finally, the advantages and limitations of each of these approaches for researching political communication are assessed. While supervised learning allows taking into account the context thanks to its ability to detect patterns of words, the researcher who uses this approach requires having data analyst skills to better refine the process. Instead, SentiStrength is more oriented to the semantic content of the terms of the message. It requires more of a competence in linguistics by the researcher. The main conclusion of this study is that both automated methods of analysis can not do without a demanding manual coding if they are to be used reliably in research.

KEY WORDS

political communication, Twitter, sentiment analysis, supervised learning, SentiStrength

1. INTRODUCCIÓN

Los mensajes publicados en Twitter constituyen un material de gran interés para detectar tendencias de opinión entre los usuarios. El hecho de que se hagan públicas opiniones, ideas y debates propicia una cierta asimilación a una conversación informal. En el contexto de la comunicación política, el análisis de contenido y los estudios cuantitativos de los mensajes de Twitter permiten identificar patrones de comportamiento entre los usuarios y puntos de inflexión en las corrientes de opinión (Jungherr, 2015).

Sin embargo, la investigación en comunicación requiere complementar los análisis cuantitativos con consideraciones de orden cualitativo. Dado el gran volumen de mensajes de Twitter que habitualmente hay que evaluar, conviene desarrollar métodos que procesen textos de forma automática con una fiabilidad de precisión aceptable. De esta forma, el investigador estaría en condiciones de cualificar mejor las opiniones y los datos extraídos de la conversación entre los usuarios. El análisis de sentimiento en Twitter surge como respuesta a esta necesidad.

El desarrollo de las técnicas de análisis de sentimiento en inglés tiene ya un cierto recorrido con estudios satisfactorios (Bravo-Marquez, Mendoza y Poblete, 2014; Da Silva, Hruschka y Hruschka, 2014; Medhat, Hassan y Korashy, 2014; Yu y Wang, 2015). Su aplicación al ámbito de la comunicación política en Twitter ha permitido alcanzar conclusiones relevantes para evaluar aspectos relacionados con la influencia (Caton, Hall y Weinhardt, 2015; Dang-Xuan, Stieglitz, Wladarsch y Neuberger, 2013) y con la agenda-setting (Guo y Vargo, 2015; Vargo, Guo, McCombs y Shaw, 2014). También estas técnicas se han aplicado a otros idiomas, como el portugués (Prata, Soares, Silva, Trevisan y Letouze, 2016). No obstante, el desarrollo de estas técnicas aplicadas al castellano se encuentra todavía en fase de maduración.

El propósito de este artículo es presentar dos técnicas de análisis de sentimiento en Twitter. En concreto, hablaremos de la familia de técnicas conocidas como aprendizaje automático (en inglés machine learning) y de la herramienta SentiStrength. El artículo se estructura del siguiente modo. En primer lugar, se introducirá el análisis de sentimiento. En segundo lugar, se presentarán las fases de la técnica basada en los algoritmos de aprendizaje automático supervisado. También se hablará de una iniciativa de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN) para potenciar el desarrollo de este tipo de análisis aplicado a mensajes escritos en castellano. A continuación, se explicará el planteamiento de la herramienta SentiStrength y los requerimientos para su uso en investigación. Se comentarán dos de los trabajos que la han empleado aplicada al castellano. Por último, se compararán las limitaciones y las posibilidades de cada una de las herramientas descritas.

2. ANÁLISIS DE SENTIMIENTO EN TWITTER

En una primera aproximación, puede decirse que analizar el sentimiento en Twitter supone asignar a cada mensaje publicado un valor relacionado con la carga emocional que transmite. En relación a esta carga emocional se pueden distinguir algunos tipos

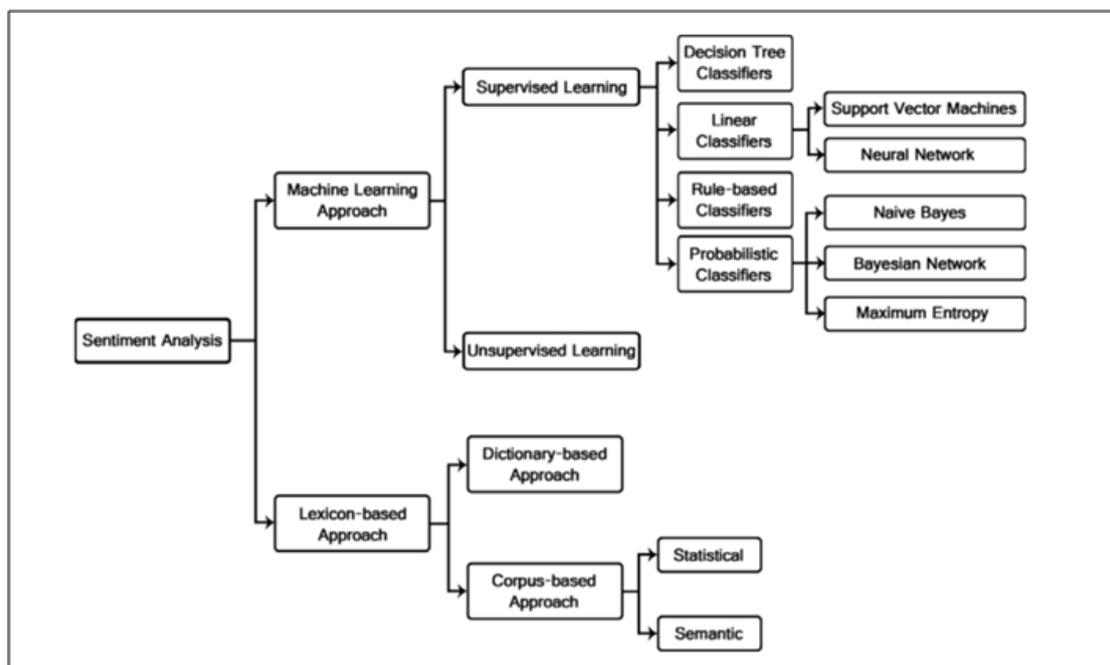
de variables diferentes (Bravo-Marquez et al., 2014):

- Polaridad: indica si el mensaje tiene un sentimiento positivo o negativo. En algunos análisis se introduce una tercera categoría para clasificar los mensajes neutros.
- Intensidad: proporciona un valor numérico en relación con la intensidad del sentimiento. Se puede distinguir entre una intensidad positiva y una intensidad negativa.
- Emoción: clasifica el texto según los distintos tipos de emociones, como puede ser la alegría, la tristeza o la ira.

La dificultad del análisis de sentimiento aplicado a grandes volúmenes de datos viene implícito en el mismo objetivo pretendido: evaluar el sentimiento o la polaridad o incluso el tipo de emoción no siempre resulta algo unívoco. Incluso cuando se realiza manualmente, el resultado puede variar según el codificador. En este punto se aprecia ya la importancia que va a tener en ambas herramientas la posibilidad de evaluar su fiabilidad.

Por tratarse de un primer acercamiento al análisis de sentimiento, se va a considerar en este artículo el nivel más sencillo: el de polaridad.

Figura 1: Tipología de las técnicas de Análisis de Sentimiento en Twitter



Fuente: Medhat et al. (2014).

Medhat et al. (2014) introducen una distinción de niveles de análisis. El sentimiento, entendido de modo general, puede ser evaluado a nivel de documento, oración o entidad. Cuando se evalúa el sentimiento de un documento, se tiene en cuenta el conjunto completo de información. El nivel de análisis de oración se limita a analizar una construcción oracional. El nivel más concreto de análisis corresponde al sentimiento

asociado a una determinada entidad, la cual podría ser un partido político, una asociación o un candidato. En este caso, quizá la evaluación del sentimiento de la frase no tiene por qué coincidir con el sentimiento asociado a la entidad estudiada. Este nivel de análisis requiere metodologías más sofisticadas y complejas. En el artículo, no se va a distinguir entre estos tres niveles, sino que se evaluará el tweet a nivel de documento independiente.

Según Medhat et al. (2014), las principales técnicas de análisis de sentimiento se dividen en dos grandes grupos: las que se basan en aprendizaje automático (machine learning approach) y las que se basan en diccionarios (lexicon-based approach). La figura 1 proporciona una visión de conjunto de estas técnicas.

Unavezbozadoslosmétodosdeanálisis,sevaaestudiarprimeroconmásdetenimiento la estrategia de los algoritmos de aprendizaje automático. Posteriormente, se tratarán las herramientas basadas en diccionarios.

3. APRENDIZAJE AUTOMÁTICO SUPERVISADO: EL CASO DE LA MÁQUINA DE VECTOR DE SOPORTE

El aprendizaje automático nació en el campo de la informática, y más en concreto, de la inteligencia artificial. Se trata de un tipo de programa informático cuyo procesamiento de datos es una suerte de aprendizaje. Dicho con otras palabras: la máquina no se programa para que responda de una determinada forma según las entradas recibidas, sino más bien para que extraiga patrones de comportamiento a partir de las entradas recibidas, y en base a dicha información aprendida o asimilada, realice la evaluación de nuevas entradas. Los algoritmos internos que constituyen la base de este aprendizaje tienen un fuerte componente estadístico y algebraico, con la consiguiente capacidad de cálculo.

Hay dos tipos de aprendizaje automático, según haya o no haya retroalimentación. Los algoritmos de aprendizaje automático no supervisado realizan el procesamiento en base únicamente a las entradas. El programa tiene la capacidad de configurarse a medida que procesa las observaciones que va recibiendo. Sería el caso, por ejemplo, de Iramuteq. Este programa realiza un análisis multidimensional de grandes volúmenes de textos, entre otros factores, a partir de la frecuencia de términos y su interconexión mutua (Ratinaud y Déjean, 2009). Los resultados responden a una lógica exploratoria, debido precisamente a su carácter no supervisado.

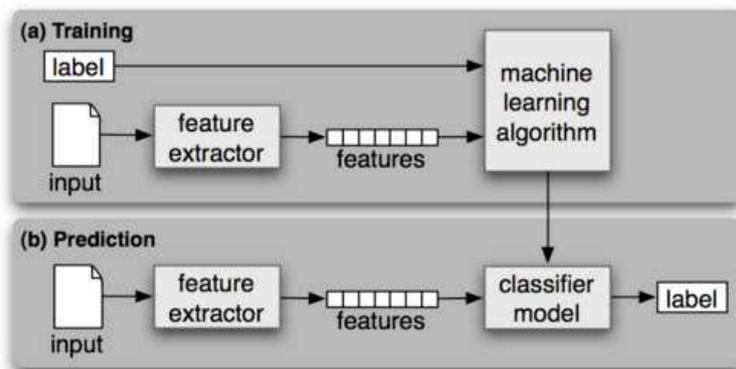
En cambio, los algoritmos de aprendizaje automático supervisado (AAS) cuentan con un corpus manualmente clasificado. Sobre este corpus particular, el algoritmo lleva a cabo dos procesos: encontrar los mejores parámetros para el algoritmo, y evaluar el nivel de fiabilidad con esos parámetros. A esta fase se le llama de aprendizaje o entrenamiento.

Un punto importante de este proceso es la elaboración del corpus para el entrenamiento. Conviene que sea lo más voluminoso posible y que sea representativo del conjunto de datos que se quiere analizar.

Debido a la posibilidad de evaluar la fiabilidad, este tipo de algoritmos supervisados son los más utilizados en el análisis de sentimiento de Twitter. Un valor mínimo de fiabilidad que podría considerarse suficiente para la investigación sería del 80%.

El siguiente paso es la fase de clasificación, la cual consiste en aplicar el algoritmo ya ajustado y entrenado a todo el volumen de datos que hay que procesar. La figura 2 ilustra este proceso.

Figura 2: Fases del AAS aplicado a la clasificación



Fuente: Bird et al. (2009).

Las familias de algoritmos supervisados pueden ser de regresión o de clasificación. Los primeros proporcionan una salida numérica mientras que los segundos asignan a la entrada una categoría proveniente de un conjunto limitado de elementos (Batrinca y Treleaven, 2015).

Para el análisis de polaridad del sentimiento en un tweet, el tipo más adecuado de AAS es el de clasificación: el algoritmo asigna una categoría al texto de entrada. Se puede discriminar entre 2 categorías (positivo / negativo) o 3 categorías (positivo / negativo / neutro).

Esta familia de técnicas depende del contexto al que se aplique. Esto significa que un algoritmo entrenado para clasificar textos de un determinado ámbito no podría ser directamente aplicado a otro ámbito. Requiere que se entrene en este nuevo contexto. Sería el caso de un algoritmo que se ha entrenado para clasificar la polaridad del sentimiento en mensajes políticos; pues bien, este algoritmo junto con los parámetros ajustados no debería ser aplicado a analizar mensajes de eventos deportivos. El error en este caso sería de orden metodológico (Saif, He, Fernandez y Alani, 2016). De todo ello se desprende la importancia clave que supone disponer de un corpus clasificado manualmente lo mejor posible.

Como puede apreciarse en la figura 2, el elemento común a la fase de aprendizaje y a la fase de clasificación es el extractor de atributos. Este módulo realiza un tratamiento de la entrada con el fin de evaluar una serie de características. Una vez extraídas, se le pasan al algoritmo, el cual tendrá una función distinta dependiendo de la fase en que se encuentre: los atributos extraídos de la entrada servirán para entrenar el algoritmo en la fase de aprendizaje, mientras que en la fase de clasificación serán utilizados para

categorizar la entrada recibida.

El preprocesamiento del tweet forma parte de este módulo, ya que los atributos no suelen extraerse directamente del texto del mensaje (Batrinca y Treleaven, 2015). Un motivo importante es subrayado por Martínez-Cámara et al. (2012) cuando apuntan la necesidad de corregir las expresiones coloquiales de los textos, dado el carácter informal de Twitter.

Un ejemplo de este preprocesamiento puede verse en el trabajo de Yu y Wang (2015). Algunos de las acciones que llevaron a cabo fueron: eliminar las URL del mensaje, pues no aportan información de sentimiento; tokenizar o extraer las palabras del tweet; pasar a minúsculas, y quitar las stopwords o palabras vacías, esto es, aquellas palabras que no aportan información de sentimiento como los artículos y las preposiciones.

Según Medhat et al. (2014), lo más habitual en el preprocesamiento es quitar las stopwords y lematizar. La unificación de las formas léxicas supone identificar términos con el mismo significado semántico, como sería el caso de una palabra en singular y otra en plural, o las diferentes formas verbales. Por ejemplo: la palabra “comerían” pasaría a ser “comer”, de tal forma que se identifica un mismo común sentido semántico en todas las formas verbales de “comer” que aparezcan en el corpus. Todas estas herramientas pertenecen al ámbito del Procesamiento de Lenguajes Naturales (PLN; en inglés Natural Language Processing, NLP).

El planteamiento general de extracción de atributos para el análisis de tweets está basado en la detección de la presencia de palabras. Como idea intuitiva, podemos afirmar que un tweet se descompone en una serie de atributos, y que cada atributo corresponde a cada una de las palabras de que consta el texto. Esta estrategia sirve para ilustrar las dimensiones del procesamiento. Un volumen considerable de tweets supondría que el elenco de posibles atributos sería enorme, tantos como palabras diferentes hubiera en los tweets del corpus de aprendizaje

Evidentemente, hay estrategias más elaboradas. Sería el caso de extraer multipalabras en lugar de unigramas (una palabra). En este caso, un atributo estaría formado por varias palabras. Lo más habitual es agrupar por dos o tres palabras. En este punto se puede tener en cuenta la estructura de la frase. Por ejemplo, se podría identificar la presencia de patrones gramaticales en los que un sustantivo estuviera seguido de otro sustantivo, o de un sustantivo seguido de una preposición y a continuación hubiera otro sustantivo (Gamallo y García, 2014).

Esta forma de extraer atributos condiciona el tipo de algoritmo a utilizar. Como entrada al algoritmo, tendremos un listado de pocos atributos por cada tweet en relación a todos los atributos posibles, es decir, la extracción conformará una matriz de datos en la que abundarán los valores nulos. Pues bien, un algoritmo muy apropiado para tratar este tipo de datos es el Support Vector Machine (SVM), o Máquina de Vectores de Soporte (MVS) (Nayak, Naik y Behera, 2015). Vapnik (1998) fue el primero en proponer este método en el área del aprendizaje estadístico. Básicamente, el algoritmo construye un enorme conjunto de hiperplanos de tal forma que maximiza los puntos de separación

entre las muestras y facilita la clasificación de las nuevas entradas.

Como hemos apuntado anteriormente, la fase de aprendizaje cuenta con un corpus clasificado manualmente y tiene un doble objetivo: ajustar el algoritmo y evaluar su fiabilidad. Por este motivo, se divide este corpus en dos conjuntos: uno más grande corresponde al corpus de desarrollo (development test) y el otro se llama corpus de prueba (test corpus). El primero se dedica propiamente al aprendizaje del algoritmo, mientras que el segundo sirve para calcular la precisión del algoritmo ya entrenado.

Hay diversas medidas para evaluar la precisión. La más sencilla y directa es la llamada accuracy (hemos decidido no traducir esta expresión puesto que otra medida más compleja es precision, y de esta forma, al mantener el término en inglés, evitamos confundirla con esta segunda medida). La accuracy evalúa el porcentaje de entradas del corpus de prueba que han sido clasificadas correctamente por el algoritmo.

Esta técnica de clasificación que combina PLN y AAS se utiliza para el análisis de sentimiento en Twitter. Ahora hemos de ver hasta qué punto proporciona resultados fiables en la clasificación de textos en castellano.

4. APLICACIÓN DEL ALGORITMO SVM AL ANÁLISIS DE SENTIMIENTO EN TWITTER EN CASTELLANO

La SEPLN convoca desde el año 2013 un Taller de Análisis de Sentimiento (TASS) en su congreso anual. Este taller se presenta como una competición para ver quién elabora el mejor algoritmo de análisis de sentimiento de acuerdo con unos objetivos previamente establecidos. Vamos a presentar los resultados del año 2015, puesto que constituyen los análisis de sentimiento en Twitter más avanzados en castellano.

El TASS 2015 estableció diversas tareas de análisis de sentimiento (García Cumbreiras, Martínez Cámara, Villena Román y García Morera, 2016). En nuestro caso, nos vamos a centrar en la más sencilla: la clasificación según la polaridad del sentimiento. Para ello, se disponía de un corpus de 68.000 tweets escritos en castellano por 150 personalidades del mundo de la política, economía, comunicación y cultura, el cual había sido redactado entre noviembre 2011 y marzo 2012 (Villena-Román, Lana-Serrano, Martínez-Cámara y González-Cristóbal, 2013). Nos referiremos a él como corpus TASS.

En este corpus, cada tweet está clasificado según una escala de seis niveles de polaridad: positivo fuerte (P+), positivo (P), neutral (NEU), negativo (N), negativo fuerte (N+) y sin sentimiento (NONE). Para la edición del TASS de 2015, el corpus se dividió en un corpus de aprendizaje (90%) y un corpus de prueba (10%). El primero se facilitó a los participantes con las etiquetas de la polaridad, para que los modelos desarrollados por cada grupo pudieran ser entrenados; mientras que las etiquetas del segundo corpus fueron retiradas con el fin de evaluar conjuntamente los resultados de todos los participantes.

A la edición del TASS del año 2015 se presentaron 17 equipos de diversas universidades

y grupos de investigación. La tarea de clasificación de la polaridad se restringió a 4 categorías: P, N, NEU y NONE. Los tres modelos con los mayores porcentaje de accuracy fueron: LIF-Run-3, con 72,6%; LIF-Run-2, 72,5%, y ELiRF-run3, 72,1% (García Cumberas et al., 2016).

Lo relevante de este taller es que se trata del estado del arte en castellano en cuanto al análisis de sentimiento mediante el procesamiento de lenguaje natural unido al AAS. Por tanto, se puede afirmar que el máximo nivel de accuracy alcanzado apenas supera el 72%. Se trata de una cifra que todavía no resulta aceptable para la investigación.

Además de este límite, conviene tener presente la complejidad de estos procesos. Vamos a describir con un poco más de detalle uno de los modelos del TASS 2015. Servirá para resaltar la elaboración que hay detrás de este tipo de análisis, pero también puede contribuir a hacerse cargo de las posibilidades que ofrecerá a medida que se mejoren estos procesos.

Los equipos participantes en el TASS podían presentar el diseño del proceso. Lamentablemente, el grupo LIF (con sus modelos LIF-Run-3 y LIF-Run-2, que fueron los que obtuvieron los mejores resultados) no publicó su informe. En cambio, el equipo ELiRF sí que dejó constancia del modelo desarrollado con el que lograron el tercer puesto en el taller (Hurtado, Pla y Buscaldi, 2015). Veamos cómo este equipo de la Universidad Politécnica de Valencia alcanzó la medida de 72,1% de accuracy.

El equipo ELiRF planteó su modelo en tres etapas. En primer lugar, se llevó a cabo un preprocesamiento de los tweets consistente en una tokenización, una lematización y un etiquetado morfosintáctico. En esta primera fase se sirvieron de la plataforma Freeling, un software para el procesamiento de lenguajes naturales desarrollado por la Universidad Politécnica de Cataluña (Padró y Stanilovsky, 2012).

A continuación, se realizó el proceso de extracción de atributos, que constituye el corazón de todo el modelo. Tuvo las siguientes fases (Pla y Hurtado, 2013):

1. Se consideraron solo unigramas de lemas con una frecuencia mínima preestablecida. Un unigrama equivale a una palabra del mensaje del tweet.
2. Los hashtag, las menciones a usuarios, los números, las fechas y los signos de puntuación fueron unificados respectivamente en una sola característica.
3. Se sustituyeron los emoticonos por su correspondiente categoría: happy, sad, tongue, wink y otros.
4. Se excluyeron los términos pertenecientes a ciertas categorías morfosintácticas poco significativas para el análisis de sentimiento.
5. Se utilizó como recurso externo varios diccionarios de polaridad.

Por último, se aplicó propiamente el algoritmo de AAS, que en este caso fue el de

máquina de vectores de soporte (SVM) tipo lineal (Hurtado et al., 2015). El resultado de todo este proceso fue una capacidad de predicción del 72,1%.

En conjunto, se puede afirmar que este tipo de técnicas que une PLN y AAS resulta altamente sofisticado, y que por ahora el resultado para el análisis de polaridad de tweets (el más simple del análisis de sentimiento) no resulta satisfactorio para su aplicación a la investigación en castellano. En cambio, hay trabajos en inglés que utilizan estas técnicas y que han obtenido niveles aceptables de accuracy. Es el caso, por ejemplo de Saif, Fernández, He y Alani (2012) y Saif, He y Alani (2012) que han alcanzado niveles entre el 80 y el 84%. Estos estudios desarrollan una nueva estrategia para extraer atributos semánticos vinculados a las entidades mencionadas en los tweets. La justificación para incorporar esta estrategia se apoya en la consistente correlación que ciertas entidades y conceptos tienden a tener con una determinada polaridad.

Uno de los elementos que utilizó el equipo ELiRF fueron los diccionarios de polaridad. Este tipo de diccionarios pueden tener un mayor protagonismo en el proceso de análisis del sentimiento de forma independiente al AAS, tal y como señalan Medhat et al. (2014). Corresponde ahora la revisión de este tipo de aproximaciones.

5. ANÁLISIS DE SENTIMIENTO BASADO EN DICCIONARIOS: EL CASO DE SENTISTRENGTH

El camino alternativo al aprendizaje automático para analizar el sentimiento de tweets se apoya en diccionarios y en algoritmos más sencillos que los vistos anteriormente.

En este contexto, un diccionario consiste en un listado de términos (pueden ser palabras o multipalabras) que tienen asociados una determinada orientación de sentimiento. De acuerdo con la distinción inicial sobre el tipo de análisis posible, esta orientación puede ser de polaridad (positiva / negativa), de intensidad de sentimiento o de tipo de emoción.

El algoritmo que evalúa la orientación de sentimiento únicamente tiene en cuenta la presencia en el texto del tweet, de términos incluidos en el diccionario. A diferencia del caso anterior, el algoritmo no se preocupa de detectar patrones sintácticos o de aprender a partir de un corpus clasificado. Su papel es más sencillo: detecta coincidencias con los diccionarios y articula un modo de evaluar el sentimiento en base al número de concurrencias encontradas.

Como puede apreciarse, el núcleo de estas técnicas se encuentra en los diccionarios de orientación del sentimiento. Este elemento depende fuertemente del idioma en que se escriba el tweet. En inglés están disponibles varios diccionarios con un alto nivel de elaboración que proporcionan información muy refinada sobre la orientación del sentimiento de los términos listados. Es el caso de SentiWordNet (Baccianella, Esuli y Sebastiani, 2010), el corpus de expresiones subjetivas Multi-perspective Question Answering (MPQA) (Wilson, Wiebey y Hoffmann, 2005) o el diccionario LIWC (Pennebaker, Mehl y Niederhoffer, 2003). En cambio, los diccionarios en castellano están más bien

orientados hacia la polarización de las palabras. Los diccionarios de Elhuyar (Saralegi y San Vicente, 2013) y el CRiSOL (Molina González, Martínez Cámara y Martín Valdivia, 2015) clasifican los términos según la polaridad positiva o negativa asociada.

La evaluación del sentimiento de una frase no radica únicamente en el contenido semántico de los términos. Hay negaciones que pueden invertir la orientación de los términos, e incluso hay veces en que coexisten términos positivos y términos negativos en la misma frase.

Un modo de superar esta limitación de los métodos basados en diccionarios sería tener en cuenta algunas reglas gramaticales y proporcionar una doble orientación del sentimiento presente en el texto. Estas premisas fueron incorporadas a la herramienta SentiStrength (Thelwall, Buckley y Paltoglou, 2012; Thelwall, Buckley, Paltoglou y Cai, 2010). Inicialmente, fue diseñado para evaluar el sentimiento en los textos publicados en MySpace. Sin embargo, la versatilidad de sus posibilidades y la facilidad de su manejo hacen que pueda ser aplicado a los textos cortos publicados en la web social, y por ello, resulta muy adecuada para los análisis de sentimiento de Twitter.

SentiStrength proporciona dos valores por cada texto analizado: uno mide la intensidad del sentimiento positivo y el otro la intensidad del sentimiento negativo. Ambos índices varían de 1 a 5. Esta doble medida se corresponde con la presencia de emociones diferentes combinadas en una misma expresión, lo cual viene sugerido por estudios empíricos del campo de la psicología (Berrios, Totterdell y Kellett, 2015).

La herramienta ofrece la posibilidad de proporcionar el resultado del análisis en otros tres formatos: binario (positivo / negativo), trinario (positivo / negativo / neutral) y en escala simple (de -4 a +4).

SentiStrength tiene como base de trabajo una serie de archivos en formato de diccionarios:

- EmotionLookUpTable.txt: lista de palabras en la que cada una tiene asociada un número que refleja su intensidad de sentimiento. El valor numérico varía entre -5 y +5.
- EmoticonLookUpTable.txt: lo mismo que el anterior pero en lugar de palabras tiene emoticonos.
- IdiomLookupTable.txt: lista de expresiones idiomáticas con su índice de sentimiento, p.ej. *a duras penas*.
- NegatingWordList.txt: lista de palabras de negación, que servirán para invertir el sentimiento de la palabra siguiente, p.ej. *jamás, nadie*.
- BoosterWordList.txt: lista de modificadores de la intensidad de sentimiento, p.ej. *muy, extremadamente*.

- SlangLookupTable.txt: traducciones de expresiones coloquiales, p.ej. q equivale a *que*.

Con este abanico de diccionarios SentiStrength permite una adaptación mayor al contexto de Twitter que la simple técnica basada en diccionarios. Conviene tener presente que una de las características del lenguaje en Twitter es su marcado carácter informal, lo cual se manifiesta en el uso de expresiones coloquiales y de emoticonos en los textos publicados. Por ello, resulta muy conveniente articular modos de evaluación de estas características.

En las primeras investigaciones donde esta herramienta fue aplicada se alcanzó un nivel de predicción del 60% para el sentimiento positivo y el 72% para el negativo (Thelwall et al., 2010). Gracias a la mejora continua del programa y de los archivos asociados, SentiStrength está siendo cada vez más utilizado en los estudios de comunicación para analizar el sentimiento de los mensajes publicados en Twitter (Bhattacharya, Srinivasan y Polgreen, 2014; Ferrara y Yang, 2015; Pfitzner, Garas y Schweitzer, 2012), incluido el ámbito de la comunicación política (Dang-Xuan et al., 2013; Guo y Vargo, 2015; Vargo et al., 2014).

Para la aplicación de SentiStrength a la investigación, la herramienta tiene la ventaja de que todos los diccionarios son editables por el usuario. Esto resulta de gran interés para la adaptación a otros idiomas. De esta forma, el usuario no requiere saber de programación, sino que su trabajo tiene un corte más lingüístico.

Además SentiStrength ofrece la posibilidad de ajustar los pesos de los términos para que se adapte mejor al contexto. Para ello, se requiere de un corpus codificado manualmente que haga la función de regla de oro. El manual de la herramienta recomienda que el volumen de este corpus sea superior a 1.000 textos y que esté validada por tres codificadores. Este proceso resulta un tanto sofisticada pues hay que asignar a cada texto dos índices de intensidad de sentimiento, uno para el positivo y el otro para el negativo.

El resultado del ajuste es una tabla nueva de los términos EmotionLookUpTable en la que el programa ha cambiado los índices de intensidad de sentimiento para que se aproxime lo máximo posible a los valores del corpus codificado manualmente. Este procesamiento interno está basado en técnicas de AAS.

La edición de los diccionarios y la posibilidad del ajuste de los pesos abren la posibilidad de aplicar el SentiStrength a otros idiomas. El objeto del siguiente apartado será comentar el uso de esta herramienta sobre textos escritos en castellano.

6. APLICACIONES DEL SENTISTRENGTH A TEXTOS EN CASTELLANO

Hemos encontrado dos estudios de comunicación política en los que se ha utilizado SentiStrength para evaluar el sentimiento de los mensajes de Twitter publicados en castellano.

Álvarez, García, Moreno y Schweitzer (2015) analizaron el papel jugado por las redes sociales para coordinar las movilizaciones en España en torno al movimiento 15M. El enfoque del estudio se dirigió al contenido semántico. Para caracterizar mejor este aspecto, combinaron el uso del diccionario LIWC (Pennebaker et al., 2003) para caracterizar las emociones transmitidas, con el uso de SentiStrength para medir la intensidad del sentimiento asociado al texto.

Para realizar la adaptación de SentiStrength al castellano, estos investigadores se valieron del corpus TASS (Villena-Román et al., 2013). Este corpus incluye tweets clasificados según temas variados, entre los que se encuentra la política y la economía.

Sin embargo, el corpus TASS tiene –a nuestro modo de ver– dos inconvenientes. En primer lugar, el paso del tiempo. Fue codificado hace más de 3 años. Este detalle influye por ejemplo en el uso de emoticonos. En aquel momento, los emoticonos estaban escritos con signos de puntuación. Sin embargo, en la actualidad la mayoría de los emoticonos utilizados en Twitter son gráficos.

El segundo inconveniente es que se trata de un corpus anotado semiautomáticamente. Esto significa que aproximadamente 7.000 tweets fueron codificados manualmente y más de 60.000 se clasificaron en base a los modelos automáticos desarrollados en las primeras ediciones del TASS. Por estos motivos, a medida que pasa el tiempo, este corpus se hace menos adecuado para el entrenamiento del SentiStrength.

El estudio de Vilares, Thelwall y Alonso (2015) tuvo un planteamiento diferente para la puesta a punto de la herramienta. En este caso, se pretendió componer explícitamente los diccionarios necesarios para la aplicación del SentiStrength a los mensajes en castellano publicados en Twitter en el ámbito de la conversación política.

Partieron de un diccionario básico que se había nutrido fundamentalmente del LIWC, y añadieron adjetivos, nombres, verbos y adverbios provenientes del estudio de Brooke et al. (2009). En este punto se puso de manifiesto un inconveniente del SentiStrength: al no realizar un procesamiento del lenguaje natural, hubo que incluir todas las formas asociadas a los términos nuevos. O sea, incluir el masculino y el femenino junto con el singular y el plural en el caso de los nombres y los adjetivos, y desarrollar las formas verbales de los infinitivos. El tamaño del diccionario principal ascendió a 26.752 entradas. Este diccionario está disponible en la siguiente dirección: <http://sentistrength.wlv.ac.uk/SpanishSentiDataDavidVilares.zip>.

La evaluación de la herramienta con estos diccionarios fue llevada a cabo con especial cuidado. Un corpus de 1.600 tweets fue manualmente codificado por tres lingüistas ajenos al proyecto. La selección del equipo se basó en que hubiera unas medidas de consistencia de intercodificación lo más altas posible. Los niveles de precisión alcanzados se midieron con un intervalo de error de +/-1. Así pues, el ajuste a los índices de sentimiento positivo fue de un 79,9% y el ajuste del sentimiento negativo fue de un 86,2% (Vilares et al., 2015).

El estudio de Vilares et al. (2015) estudio proporciona una base inicial de trabajo

bastante fiable para aplicar la herramienta del SentiStrength a los mensajes publicados en Twitter.

7. DISCUSIÓN

Medhat et al. (2014) habían distinguido dos planteamientos para evaluar el sentimiento en los mensajes de Twitter: uno basado en el procesamiento del lenguaje natural unido a técnicas de AAS, y otro que utiliza diccionarios de polaridad o de sentimiento. Sin embargo, tras ver con más de detalle ejemplos concretos de estos planteamientos, se ha puesto de manifiesto que ambos enfoques comparten elementos. Los procesos de AAS pueden servirse de diccionarios de polaridad para extraer atributos de los tweets; mientras que el SentiStrength, que está basado en el uso de diccionarios, utiliza técnicas de AAS para ajustarse al contexto. Una primera consecuencia emerge ya del presente estudio: no conviene obviar uno a favor de otro, puesto que las técnicas básicas se utilizan en ambos enfoques.

En estos dos planteamientos se requiere un corpus codificado manualmente, si bien la necesidad de este corpus difiere ligeramente. En el caso de los procesos de AAS, el corpus tiene un papel clave. Conviene que sea suficientemente representativo y voluminoso, y, sobre todo, que cuente con una validación completa de tres codificadores. Para el uso de SentiStrength en investigación, se necesita un corpus codificado igualmente con la validación de tres codificadores. La dificultad en este caso estriba en la exigencia de la codificación. Así como en el caso del AAS las categorías pueden ser simplemente de polaridad (positiva / negativa), en el caso del SentiStrength se han de señalar dos índices que reflejen la intensidad positiva y negativa de cada tweet de 1 a 5.

Una diferencia entre ambos planteamientos radica en la relación con el contexto. Los métodos que utilizan el AAS dependen del contexto, dado que necesitan un corpus de entrenamiento. Conviene que este corpus sea un extracto del corpus general, de tal forma que asegure una coherencia en la clasificación. En cambio, los diccionarios tienen una polaridad asociada a las palabras que permite usar estos métodos en diversos contextos.

Ahora bien, la evaluación del sentimiento puede devenir un tanto rígida en los métodos basados en diccionarios, puesto que depende exclusivamente del contenido semántico de los términos. En cambio, el AAS permite una mayor adaptación, dado que el valor añadido que aporta este tipo de procesos es la detección de patrones de palabras que transmiten un sentimiento particular. Lo que sí que conviene señalar en ambos métodos es la limitación para detectar correctamente expresiones irónicas. En este sentido, el SentiStrength intenta superar esta dificultad al proporcionar la posibilidad de interpretar este tipo de expresiones, dado que incluye un diccionario específico para ello.

En general, se puede afirmar que los métodos de análisis de sentimiento que emplean el AAS ponen el énfasis en los procesos, mientras que las herramientas como el SentiStrength lo ponen en los diccionarios. Este enfoque diferente determina la habilidad requerida por parte del investigador: el uso del AAS presupone una competencia en la

disciplina emergente de análisis de datos, mientras que el SentiStrength reclama más bien una competencia de orden lingüístico. La tabla 1 recoge las principales diferencias entre ambos enfoques, particularizando para los casos vistos en este artículo: el PLN con AAS-MVS y el SentiStrength.

Tabla 1: Comparación entre los métodos de análisis de sentimiento en Twitter

	PLN + AAS-MVS	SentiStrength
Base de trabajo	Procesos	Diccionarios
Habilidad del Investigador	Analista de Datos	Lingüística
Codificación manual	Corpus de Entrenamiento	Regla de Oro para el ajuste de pesos
Evaluación	Accuracy de la Clasificación	Precisión de la Validación
Adaptación	Flexible por el Entrenamiento	Rígida por los Diccionarios
Contexto	Dependiente	Independiente
Valor aportado	Detecta patrones no evidentes	Sencillez de uso
Dificultad	Alta	Moderada - Alta

Fuente: elaboración propia

8. CONCLUSIONES

El análisis de sentimiento en Twitter aporta información relevante para el análisis de la comunicación que se da en este espacio de Internet. El hecho de tener textos cortos en cada mensaje hace razonable intentar desarrollar métodos que evalúen de forma automática la carga de sentimiento transmitida en cada tweet. El desarrollo de estos procesos aplicados a los mensajes en castellano está siendo objeto de creciente interés por parte de los investigadores en comunicación y del sector de la ingeniería computacional.

Captar el sentimiento de una frase depende de la construcción sintáctica y especialmente del contenido semántico. A partir de esta doble fuente se están desarrollando técnicas diferentes: una toma como base el AAS unido al procesamiento del lenguaje natural, y la otra emplea diccionarios con información de polaridad de sentimiento. Ambos planteamientos metodológicos no están tan separados como podría parecer, pues comparten técnicas características de uno y otro.

Sin embargo, el estudio en detalle de una aplicación de cada uno de estos enfoques revela una diferencia de competencias requeridas en el investigador, así como una evidente diferencia de dificultad en su uso. El análisis de datos aplicado a Twitter abre posibilidades prometedoras al constatar los logros en inglés. De igual forma, la puesta a disposición de diccionarios de polaridad en castellano cada vez mejores, facilitará que estos sistemas puedan extenderse cada vez más.

El presente estudio revela también algunas líneas de investigación en el campo del análisis de sentimiento en Twitter. Los niveles alcanzados de precisión en las mejores prácticas publicadas que utilizan el AAS son todavía insuficientes para su aplicación a la investigación. No obstante, cabe hacer una observación: estas prácticas utilizaron un corpus para su aprendizaje en el que se acusa ya el paso del tiempo. Bien es cierto que alcanzar el volumen de tweets del corpus TASS no es nada sencillo, pero si se desea aplicar los métodos de AAS convendrá codificar un corpus manualmente que sea suficientemente grande.

En el caso de SentiStrength, dada su facilidad de uso, se cuenta con el corpus desarrollado por Vilares et al. (2015) que se encuentra a disposición de los investigadores. Sin embargo, conviene realizar antes una evaluación de la precisión dado que esta herramienta se encuentra todavía en fase de maduración en castellano. Por este motivo, las pautas proporcionadas por su desarrollador para el ajuste de pesos constituyen un paso muy recomendable para asegurar que la herramienta proporciona un análisis fiable del sentimiento. Esta tarea no es baladí, pero seguramente un buen proceso de ajuste proporcionará posibilidades para la investigación en comunicación de gran interés para conocer mejor los procesos de influencia en Twitter.

AGRADECIMIENTOS

Este trabajo ha sido posible gracias a Emilio Giner Hernández por introducir al autor en el análisis automatizado de sentimiento, y a Lorena Cano Orón por sugerir redactarlo. La presente investigación forma parte del proyecto I+D “Los flujos de comunicación en los procesos de movilización política: medios, blogs y líderes de opinión” financiado por el Ministerio de Economía y Competitividad (referencia CSO2013-43960-R).

REFERENCIAS

Álvarez, R.; García, D.; Moreno, Y. & Schweitzer, F. (2015): “Sentiment cascades in the 15M movement”, *EPJ Data Science*, vol. 4, no 6. <http://dx.doi.org/10.1140/epjds/s13688-015-0042-4>

Baccianella, S.; Esuli, A. & Sebastiani, F. (2010): “SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining”. Comunicación presentada en el *Seventh Conference on International Language Resources and Evaluation*. Valetta, Malta, 19-21 de mayo. Disponible en Internet: <http://nmis.isti.cnr.it/sebastiani/Publications/LREC10.pdf> [Consulta: 12/10/2016].

Batrinca, B. & Treleaven, P. C. (2015): “Social media analytics: a survey of techniques, tools and platforms”, *AI and Society*, vol. 30, no 1, pp. 89–116. <http://dx.doi.org/10.1007/s00146-014-0549-4>

Berrios, R.; Totterdell, P. & Kellett, S. (2015): “Eliciting mixed emotions: a meta-analysis comparing models, types, and measures”, *Frontiers in Psychology*, vol. 6, no 428. <http://dx.doi.org/10.3389/fpsyg.2015.00428>

Bhattacharya, S.; Srinivasan, P. & Polgreen, P. (2014) “Engagement with Health Agencies on Twitter”, *PLoS ONE*, vol. 9, no 11, pp. e112235. <http://dx.doi.org/10.1371/journal.pone.0112235>

Bird, S.; Klein, E. & Loper, E. (2009): *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. Sebastopol, CA: O’Reilly Media.

Bravo-Marquez, F.; Mendoza, M. & Poblete, B. (2014): “Meta-level sentiment models for big social data analysis”, *Knowledge-Based Systems*, vol. 69, no 1, pp. 86–99. <http://dx.doi.org/10.1016/j.knosys.2014.05.016>

Brooke, J.; Tofiloski, M. & Taboada, M. (2009): "Cross-Linguistic Sentiment Analysis: From English to Spanish". Comunicación presentada en el *International Conference on Recent Advances in Natural Language Processing (RANLP)*. Borovets, Bulgaria, 14-16 de septiembre. Disponible en Internet: <http://aclweb.org/anthology/R/R09/R09-1010.pdf> [Consulta: 12/10/2016].

Caton, S.; Hall, M. & Weinhardt, C. (2015): "How do politicians use Facebook? An applied Social Observatory", *Big Data & Society*, vol. 2, no 2. <http://dx.doi.org/10.1177/2053951715612822>

Da Silva, N. F. F.; Hruschka, E. R. & Hruschka, E. R. (2014): "Tweet sentiment analysis with classifier ensembles", *Decision Support Systems*, vol. 66, pp. 170–179. <http://dx.doi.org/10.1016/j.dss.2014.07.003>

Dang-Xuan, L.; Stieglitz, S.; Wladarsch, J. & Neuberger, C. (2013): "An Investigation of Influentials and the Role of Sentiment in Political Communication on Twitter During Election Periods", *Information, Communication & Society*, vol. 16, no 5, pp. 795–825. <http://dx.doi.org/10.1080/1369118X.2013.783608>

Ferrara, E. & Yang, Z. (2015): "Measuring Emotional Contagion in Social Media", *PLoS ONE*, vol. 10, no 11, pp. e0142390. <http://dx.doi.org/10.1371/journal.pone.0142390>

Gamallo, P. y García, M. (2014): "Citius: A Naive-Bayes Strategy for Sentiment Analysis on English Tweets". Comunicación presentada en el *8th International Workshop on Semantic Evaluation (SemEval 2014)*. Dublin, 23-24 de agosto.

García Cumbreñas, M. Á.; Martínez Cámara, E.; Villena Román, J. & García Morera, J. (2016): "TASS 2015 - The evolution of the Spanish opinion mining systems", *Revista de Procesamiento Del Lenguaje Natural*, vol. 56, pp. 33–40. Disponible en Internet: <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/5284/3078> [Consulta: 10/10/2016].

Guo, L. & Vargo, C. (2015): "The Power of Message Networks: A Big-Data Analysis of the Network Agenda Setting Model and Issue Ownership", *Mass Communication and Society*, vol. 18, no 5, pp. 557–576. <http://dx.doi.org/10.1080/15205436.2015.1045300>

Hurtado, L.-F.; Pla, F. & Buscaldi, D. (2015): "ELiRF-UPV en TASS 2015: Análisis de Sentimientos en Twitter". Comunicación presentada en el *TASS: Workshop on Sentiment Analysis at SEPLN*. Alicante, 15 de septiembre. Disponible en Internet: http://ceur-ws.org/Vol-1397/elirf_upv.pdf [Consulta: 6/10/2016].

Jungherr, A. (2015): *Analyzing Political Communication with Digital Trace Data: The Role of Twitter Messages in Social Science Research*. Cham: Springer International Publishing Switzerland.

Martínez-Cámara, E.; Martín-Valdivia, M. T.; Ureña-López, L. A. & Montejo-Ráez, A. (2012): "Sentiment analysis in Twitter", *Natural Language Engineering*, vol. 20, no 1, pp.

1–28. <http://dx.doi.org/10.1017/S1351324912000332>

Medhat, W.; Hassan, A. & Korashy, H. (2014): "Sentiment analysis algorithms and applications: A survey", *Ain Shams Engineering Journal*, vol. 5, no 4, pp. 1093–1113. <http://dx.doi.org/10.1016/j.asej.2014.04.011>

Molina González, M. D.; Martínez Cámara, E. & Martín Valdivia, M. T. (2015): "CRiSOL: Base de conocimiento de opiniones para el español", *Procesamiento Del Lenguaje Natural*, vol. 55, pp. 143–150. Disponible en Internet: <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/5226> [Consulta: 12/10/2016].

Nayak, J.; Naik, B. & Behera, H. S. (2015): "A Comprehensive Survey on Support Vector Machine in Data Mining Tasks: Applications & Challenges", *International Journal of Database Theory and Application*, vol. 8, no 1, pp. 169–186. <http://dx.doi.org/10.14257/ijdta.2015.8.1.18>

Padró, L. & Stanilovsky, E. (2012): "FreeLing 3.0: Towards Wider Multilinguality". Comunicación presentada en el *Language Resources and Evaluation Conference (LREC 2012) ELRA*. Estambul, Turquía, 22-25 de mayo. Disponible en Internet: http://www.lrec-conf.org/proceedings/lrec2012/pdf/430_Paper.pdf [Consulta: 25/05/2016].

Pennebaker, J. W.; Mehl, M. R. & Niederhoffer, K. G. (2003): "Psychological Aspects of Natural Language Use: Our Words, Our Selves", *Annual Review of Psychology*, vol. 54, pp. 547–577. <http://dx.doi.org/10.1146/annurev.psych.54.101601.145041>

Pfitzner, R.; Garas, A. & Schweitzer, F. (2012): "Emotional Divergence Influences Information Spreading in Twitter". Comunicación presentada en el *Sixth International AAAI Conference on Weblogs and Social Media*. Dublín, 4-7 de junio. Disponible en Internet: <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4596/5053> [Consulta: 15/10/2016].

Pla, F. & Hurtado, L.-F. (2013): "ELiRF-UPV en TASS 2013: Análisis de Sentimientos en Twitter". Comunicación presentada en el *TASS workshop at SEPLN 2013. IV Congreso Español de Informática*. Madrid, 18-20 de septiembre. Disponible en Internet: http://users.dsic.upv.es/~lhurtado/papers/pdfs/2013_pla13_tass.pdf [Consulta: 6/10/2016].

Prata, D. N.; Soares, K. P.; Silva, M. A.; Trevisan, D. Q. & Letouze, P. (2016): "Social Data Analysis of Brazilian's Mood from Twitter", *International Journal of Social Science and Humanity*, vol. 6, no 3, pp. 179–183. <http://dx.doi.org/10.7763/IJSSH.2016.V6.640>

Ratinaud, P. & Déjean, S. (2009): "Iramuteq. Implémentation de la méthode Alceste d'analyse de texte dans un logiciel libre". Comunicación presentada en el *Modélisation appliquée aux sciences humaines et sociales (MASHS 2009)*. Toulouse, 10 de junio.

Saif, H.; Fernandez, M.; He, Y. & Alani, H. (2012): "Semantic sentiment analysis of Twitter". Comunicación presentada en el *1th International Conference on the Semantic Web*. Boston, MA, 11-15 de noviembre. http://dx.doi.org/10.1007/978-3-642-35176-1_32

Saif, H.; He, Y. & Alani, H. (2012): "Alleviating data sparsity for Twitter sentiment analysis". Comunicación presentada en el *2nd Workshop on Making Sense of Microposts (#MSM2012) at the 21st International Conference on the World Wide Web (WWW'12)*. Lyon, Francia, 16 de abril.

Saif, H.; He, Y.; Fernandez, M. & Alani, H. (2016): "Contextual semantics for sentiment analysis of Twitter", *Information Processing and Management*, vol. 52, no 1, pp. 5–19. <http://dx.doi.org/10.1016/j.ipm.2015.01.005>

Saralegi, X. & San Vicente, I. (2013): "Elhuyar at TASS 2013". Comunicación presentada en el *XXIX Congreso de la Sociedad Española de Procesamiento de Lenguaje Natural*. Madrid, 18-20 de septiembre. Disponible en Internet: <http://www.sepln.org/workshops/tass/2013/papers/tass2013-submission3-Elhuyar.pdf> [Consulta: 6/10/2016].

Thelwall, M.; Buckley, K. & Paltoglou, G. (2012): "Sentiment Strength Detection for the Social Web", *Journal of the American Society for Information Science and Technology*, vol. 63, no 1, pp. 163–173. <http://dx.doi.org/10.1002/asi.21662>

Thelwall, M.; Buckley, K.; Paltoglou, G. & Cai, D. (2010): "Sentiment Strength Detection in Short Informal Text", *Journal of the American Society for Information Science and Technology*, vol. 61, no 12, pp. 2544–2558. <http://dx.doi.org/10.1002/asi.21416>

Vapnik, V. V. (1998): *Statistical Learning Theory*. New York: John Wiley & Sons.

Vargo, C. J.; Guo, L.; McCombs, M. & Shaw, D. L. (2014): "Network Issue Agendas on Twitter During the 2012 U.S. Presidential Election", *Journal of Communication*, vol. 64, pp. 296–316. <http://dx.doi.org/10.1111/jcom.12089>

Vilares, D.; Thelwall, M. & Alonso, M. A. (2015): "The megaphone of the people? Spanish SentiStrength for real-time analysis of political tweets", *Journal of Information Science*, vol. 41, no 6, pp. 799–813. <http://dx.doi.org/10.1177/0165551515598926>

Villena-Román, J.; Lana-Serrano, S.; Martínez-Cámara, E. & González-Cristóbal, J. C. (2013): "TASS - Workshop on Sentiment Analysis at SEPLN", *Revista de Procesamiento Del Lenguaje Natural*, vol. 50, pp. 37–44. Disponible en Internet: <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/4657/2759> [Consulta: 6/10/2016]

Wilson, T.; Wiebe, J. & Hoffmann, P. (2005): "Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis". Comunicación presentada en el *Conference on Human Language Technology and Empirical Methods in Natural Language Processing*. Vancouver. Disponible en Internet: <https://people.cs.pitt.edu/~wiebe/pubs/papers/emnlp05polarity.pdf> [Consulta: 12/10/2016].

Yu, Y. & Wang, X. (2015): "World Cup 2014 in the Twitter World: A big data analysis of sentiments in U.S. sports fans' tweets", *Computers in Human Behavior*, vol. 48, pp. 392–400. <http://dx.doi.org/10.1016/j.chb.2015.01.075>