

Selección de características para determinar la polaridad de tuits en idioma español a nivel global

Armando Reyes Correa, José Luis Tapia Fabela, Yulia Ledeneva,
René Arnulfo García-Hernández, Rafael Cruz Reyes

Universidad Autónoma del Estado de México,
Toluca, México

armando_reyesc@cbtis162.edu.mx,
joseluis.fabela@gmail.com, yledeneva@yahoo.com,
rearnulfo@hotmail.com, rcruzrey@gmail.com

Resumen. En la actualidad, el internet ha transformado la forma en como las personas se conducen en sus vidas personales y empresariales. Estos cambios son impulsados principalmente por los contenidos generados por los usuarios, mediante opiniones expresadas en forma de texto en las redes sociales. De acuerdo a la Sociedad Española del Lenguaje Natural (SEPLN), el idioma español ocupa el segundo lugar a nivel mundial como lengua materna y como segundo idioma más utilizado en internet, sin embargo, su importancia como lenguaje no corresponde con el mismo nivel de investigación del que es objeto, ya que la mayoría de estas investigaciones se realiza para el idioma inglés. La SEPLN con el fin de promover la investigación para el descubrimiento de nuevos algoritmos y técnicas en análisis de sentimientos en twitter en español, convoca a participar en el taller denominado Análisis de Sentimientos (TASS), que, en su última edición, se invita a participar en dos tareas, Análisis de Sentimientos a Nivel Global y Análisis de Sentimientos a Nivel Aspecto. Este trabajo, se enfoca al problema de la primera tarea. Utilizando los textos que componen el corpus en español otorgado por el TASS 2015, y seleccionando los tres clasificadores más utilizados por el estado del arte Naive Bayes, MSV y J48 mediante el software Weka, en este trabajo se presentan dos modelos, el primero sin ningún tipo de pre-procesamiento y el segundo utilizando características léxicas de los textos, para clasificarlos en seis y cuatro categorías tal como lo define la tarea del concurso.

Palabras clave: Análisis de sentimientos, características léxicas, Naive Bayes, MSV, J48.

Features Selection to Define the Polarity of Tweets in Spanish at Global Level

Abstract. Nowadays, Internet has transformed the way in how people conducted in their personal lives and business. This change are mainly proposed by the user contents generated, in opinions expressed by text form in social media. According to Spanish Society of Natural Language Processing (Sociedad Española de Lenguaje Natural, SEPLN), the Spanish language has the second place in importance in the world as mother language and as second language in Internet, however, this fact, does not correspond with the same level of research, due to the most research is for English Language. The aim of SEPLN, through the Workshop of Sentiment Analysis (SA) denominated TASS is to provide a forum where the newest research works in algorithms and in SA techniques in social media, are showed and discussed by scientific and business communities. The last edition of TASS convenes to participate in two tasks, Sentiment Analysis at Global Level, and Sentiment Analysis at Aspect Level. This work is focused in the first task, using the corpus provided by the TASS 2015 workshop, we build two models in Naïve Bayes, Support Vector Machines and J48 classifiers in Weka software, the first one without preprocessing and the second using the main lexical features of the corpus to classify in six and four categories.

Keywords. Sentiment analysis, lexical features, Naïve Bayes, SVM, J48.

1. Introducción

Con el origen de la Web 2.0, Internet contiene grandes cantidades de información generada por el usuario en un ilimitado número de temas. Muchas entidades tales como corporaciones o grupos políticos tratan de obtener conocimiento a través de las opiniones expresadas por los usuarios. Las plataformas sociales tales como Facebook o Twitter han probado ser exitosas para estas tareas, debido al alto volumen de mensajes en tiempo real que se generan y el gran número de usuarios que las utilizan todos los días [1].

Actualmente la gran cantidad de datos almacenados en medios electrónicos, unido al desarrollo tecnológico de las computadoras agrupados bajo el término conocido como “data mining”, tiene como objetivo extraer información y conocimiento útil para aplicarlo en cualquier área productiva apoyando la toma de decisiones [2]. El área de estudio del procesamiento de lenguaje natural encargado de analizar y clasificar textos en polaridades positivas, negativas o neutrales se le denomina análisis de sentimientos, conocido también como minería de opiniones, análisis de subjetividad y orientación de sentimientos [9]. La minería de opiniones [22], en un sentido amplio, se define como el estudio computacional de opiniones, sentimientos y emociones expresadas en el texto.

Detectar sentimientos es considerada una tarea difícil, ya que el problema implica el conocimiento del entorno y contexto donde se ejecuta la opinión, el cual es muy amplio y complejo [23]. Formalmente se dice que una opinión de una característica c conlleva un sentimiento asociado, el usuario que emite la opinión es conocido como el emisor de la opinión. De esta forma, una opinión es definida como una quintupla $(o_j, f_{jk}, oo_{ijkl}, h_i, t_i)$ [8] donde:

o_j - es el objeto de la opinión,

f_{jk} - es una característica del objeto sobre el que se expresa una opinión. Cuando no se detecta ninguna característica, se interpreta como una opinión general, como característica del objeto,

oo_{ijkl} - es la polaridad del sentimiento de la opinión sobre la característica f_{jk} del objeto o_j - positivo, negativo o neutral,

h_i - es el emisor de la opinión,

t_i - es el tiempo en que la opinión es expresada por h_i .

La Sociedad Española del Procesamiento del Lenguaje Natural (SEPLN), una asociación científica, sin fines de lucro que promueve la investigación de todo tipo de actividades relacionadas con el estudio del procesamiento del lenguaje natural en español [7], organiza a partir del año 2012, un taller en Análisis de Sentimientos enfocado en el idioma español sobre textos extraídos de la red social Twitter. Twitter es una plataforma de microblogging donde los usuarios publican mensajes, opiniones y comentarios cuyos contenidos van del rango de los sentimientos personales a publicaciones generales. Las publicaciones en Twitter se conocen como tuits. La característica principal de los tuits es que la longitud máxima del texto es de 140 caracteres [23].

En la última edición de este taller, se organizan dos tareas las cuales son: Análisis de Sentimientos a Nivel Global y Análisis de Sentimientos a Nivel Aspecto. El trabajo presentado, se enfoca en la primera tarea del taller, que consiste en determinar la polaridad global de los mensajes en twitter categorizándolos en seis y cuatro categorías: (P+, P, NEU, N, N+, NONE) y (P, NEU, N, NONE).

Este artículo resume la primera aproximación experimental para determinar la polaridad de tuits en el idioma español, usando el corpus de TASS 2015 y los clasificadores de Weka: J48, Naive Bayes y Máquina de Soporte Vectorial. Las características léxicas extraídas son incluidas como datos de entrada en los clasificadores mencionados, el propósito de este trabajo es analizar las entradas y salidas requeridas por los clasificadores, además de identificar qué ajustes en sus parámetros se logran los mejores resultados.

El resto del artículo se compone de la siguiente manera: Sección 2 resume algunos trabajos publicados en TASS, en la sección 3 se describe la tarea y se analiza el corpus proporcionado, en la sección 4 se explica la parte de experimentación y resultados obtenidos, en la sección 5 se muestra una comparación de los resultados obtenidos con el estado del arte, en la sección 6 se presentan conclusiones y trabajo futuro.

2. Estado del arte

Existen dos enfoques para clasificar comentarios y opiniones en positivos, negativos o neutrales: algoritmos de aprendizaje supervisado y no supervisado. Los algoritmos de aprendizaje supervisado, son utilizados en los problemas donde se conoce a priori el número de clases y miembros representativos de cada clase. La tarea de clasificación de sentimientos, puede ser formulada como un problema de aprendizaje supervisado. Los clasificadores más utilizados en análisis de sentimientos son Naive Bayes (NB), Máquinas de Soporte Vectorial (MSV) y Máxima Entropía. En muchos casos los algoritmos tipo MSV han mostrado una mejora notable sobre Naive Bayes [3].

En el estado del arte estudiado, se ha encontrado que en ediciones pasadas del TASS, el problema de categorizar en seis y cuatro etiquetas se ha abordado mediante el enfoque de aprendizaje supervisado y no supervisado. En el enfoque no supervisado se encuentra el trabajo desarrollado por el equipo SINAI-ESMA en su participación en la edición de TASS 2014 [5], basado en el uso de léxico de opinión y aplicación de una heurística sintáctica, logrando una exactitud en 4 categorías de 0.6456 y en 6 categorías de 0.5360. El equipo del Instituto Politécnico Nacional en el artículo “Análisis de Sentimiento sobre textos en español basado en aproximaciones semánticas con reglas lingüísticas” [6], explica que la clasificación de la polaridad la realiza de acuerdo a un diccionario de orientación semántica, mostrando una exactitud de 0.7648 en 3 categorías y de 0.6261 en 5 categorías.

Analizando los cuatro primeros lugares de las ediciones del TASS 2012 a 2015, se encontró que todos ellos utilizan el enfoque de aprendizaje supervisado. En 2012, el equipo Elhuyar Fundazioa, utiliza un lexicón de polaridad para español a partir de uno en idioma inglés mediante el modelo de unigramas, utilizando como características emoticones negativos y positivos, así como intersecciones positivas y negativas para entrenar un clasificador tipo MSV alcanzando una exactitud de 0.702 en la clasificación de 4 categorías y de 0.641 en seis categorías [12]. En 2013 el equipo DLSI-UA presentó resultados de 0.616, lo que lo ubicó como el primer lugar del concurso, en una clasificación de 6 categorías, utilizando los modelos bigramas y skipgramas apoyándose con un lexicón de sentimientos, para entrenar un clasificador tipo MSV [13]. En ese mismo año, el equipo Elhuyar Fundazioa, segundo lugar de esa edición presentó una metodología creando una lista de vocabulario coloquial a partir de un diccionario de expresiones coloquiales, seleccionando como características los lemas que corresponden a las palabras incluidas en el lexicón de polaridad. El equipo entrena una MSV alcanzando resultados de 0.686 en cuatro categorías y 0.601 en 6 categorías [14]. Para la edición del 2014, el equipo Elhuyar, construye un modelo en Weka, utilizando una MSV, combinando la información extraída de léxicos de polaridad con características lingüísticas. Utiliza un modelo basado en ngramas basados en patrones sintácticos [N+Adj] y [Verb+Noun], con un umbral de 3 ocurrencias para tomar en cuenta los ngramas como características, así mismo, hace uso de los signos de puntuación, tratamiento de la negación y combinación de lexicones. Presenta como resultados 0.6990 para la clasificación en 4 categorías, mientras que para la de seis categorías obtiene un 0.6100 ubicándolo en el segundo lugar [15]. El primer lugar, el equipo EliRF-UPV utiliza el modelo de ngramas de palabras y de lemas, tomando las

características tf-idf de los lemas y de las palabras que aparecen en el tuit, más el número de lemas y palabras que aparecen en el diccionario de lemas y palabra respectivamente. En una tercera aproximación, ejecuta una votación entre 6 sistemas de 1, 2 y 3 gramas de palabras y lemas. Los resultados más altos publicados son 0.6432 para la clasificación de 6 categorías y 0.7089 para la de cuatro categorías. En la última edición del TASS, en el año 2015, el equipo LIF, desarrolla una arquitectura de dos niveles, reproduciendo en el primer nivel 5 sistemas con base en MSV, Redes Neuronales Convolucionales y la implementación del enfoque propuesto por [4]. Este consiste en extraer características doc2vec de los tuits y utilizarlas como entradas en un clasificador del tipo MSV. En el segundo nivel, entrena una MSV, fusionando los resultados del primer nivel para ingresarlos como características. LIF consiguió el primer lugar de esa edición del TASS obteniendo una precisión de 0.6720 para seis categorías y 0.7260 en la clasificación de 4 categorías [9]. El equipo ELiRF-UPV, logró el segundo lugar representando al texto como: votación simple de palabras, lemas y ngramas, además, de utilizar como clasificador una MSV con una precisión de 0.6730 en la clasificación de 6 categorías y 0.7250 en 4 categorías [10].

Otros trabajos analizados es el Sidorov [23], quienes en 2014 desarrollan un lexicón de emociones en español, catalogado como un recurso novedoso para el análisis de emociones en textos, etiquetado con probabilidades que expresan una de las seis emociones básicas. En este trabajo, exploran diferentes configuraciones (tamaño de n-gramas, tamaño del corpus, número de clases de sentimientos, corpus balanceados contra no balanceados, varios dominios) para determinar cómo afectan la precisión de diferentes algoritmos de aprendizaje, utilizando la API de Weka para los clasificadores Naive Bayes, Árboles de Decisión y MSV. Los datos de estudio se componen de tuits en idioma español, recopilado por los autores sobre entidades predefinidas de marcas de teléfonos celulares. Se recopilaron 32,000 tuits y alrededor de 8,000 tuits fueron etiquetados manualmente en una de cuatro categorías, positivo, negativo, neutral o informativo.

En sus resultados concluyen que un conjunto de 3,000 instancias es suficiente como conjunto de entrenamiento para un tema seleccionado; se indica que el unigrama es la mejor longitud utilizada como característica; el número de clases afecta el rendimiento de los clasificadores, reduciendo el número de clases se aumenta la precisión del clasificador; utilizando un corpus no balanceado se logra mejor precisión alcanzando un 0.8580 con un clasificador MSV. Se comprueba que el dominio es parte importante de la precisión alcanzada, ya que, en un experimento, se entrenó los clasificadores utilizando datos del dominio de teléfonos celulares y las pruebas se realizaron con datos del mismo dominio y con un corpus de tuits en el dominio del tema político. Los resultados muestran que el cambio de dominio en la fase de prueba afecta negativamente la precisión pasando de un 0.858 en el mismo dominio en una MSV a un 0.280 en el dominio del tema político utilizando el mismo clasificador. En el trabajo desarrollado por Sanzón [21], presentan dos modelos para tratar el corpus otorgado por SemEval 2014, el primero basado en características léxico - sintácticas y el segundo modelo mediante la representación de grafos, selección de características y representación vectorial. Para el primer modelo, se desarrollan dos diccionarios de forma manual sobre emoticones y otro de siglas empleadas en redes sociales,

posteriormente se realiza un pre procesamiento en los datos de entrenamiento y prueba. Finalmente se crean los modelos utilizando los clasificadores Naive Bayes y MSV. Para el segundo modelo, se utiliza el mismo procedimiento de normalización y pre proceso que el primer modelo, los textos se representan mediante grafos de co-ocurrencia no dirigidos utilizando ventanas de tamaño 2 y 3. Los resultados en porcentaje de precisión obtenida en los dos modelos son: En el primer modelo utilizando Naive Bayes 0.4726 y con MSV 0.5658. El modelo de grafos utilizando una MSV, para una ventana de tamaño 2, el mejor resultado obtenido es de 0.4734 y con ventana de tamaño 3 es de 0.4564.

3. Descripción de la tarea y análisis del corpus

Para el TASS 2015, se repite la tarea de ediciones anteriores, que consiste en realizar una clasificación automática para determinar la polaridad global de cada mensaje, validando los modelos sobre el corpus general de prueba. Hay dos evaluaciones diferentes: una basada en 6 etiquetas de polaridad diferente (P+, P, NEU, N, N+, NONE) y otra basada en sólo 4 etiquetas (P, NEU, N, NONE) [7].

El corpus proporcionado por TASS consta de un conjunto de entrenamiento compuesto por 7,219 tuits etiquetados con la polaridad correspondiente a 6 etiquetas. Se muestra la distribución de tuits por su polaridad en el conjunto de entrenamiento

Tabla 1. Distribución de tweets en el corpus de entrenamiento.

Categorías	No. de Tuits	%
+P	1652	22.88%
P	1232	17.07%
NEU	670	9.28%
N	1335	18.49%
+N	847	11.73%
NONE	1483	20.54%

Se cuenta también con un corpus general utilizado como conjunto de prueba compuesto por 60,798 tuits.

La evaluación de los sistemas desarrollados, define el TASS en su sitio web, será utilizando la métrica de Exactitud (Accuracy) [11], la cual evalúa la polaridad correcta asignada a los tuits de acuerdo al estándar de oro. El corpus de prueba, es utilizado para evaluar la exactitud del aprendizaje del modelo, apoyado por la disposición del conjunto de prueba etiquetado. La exactitud de un modelo de clasificación sobre un conjunto de prueba es definida como [16]:

$$Exactitud = \frac{\text{Número correcto de clasificaciones}}{\text{Número total de casos de prueba}}$$

La matriz de confusión generada, será utilizada para evaluar precisión, recuerdo y medida - F1 para cada categoría individual. La metodología utilizada como primer aproximación se representa en la figura 1.

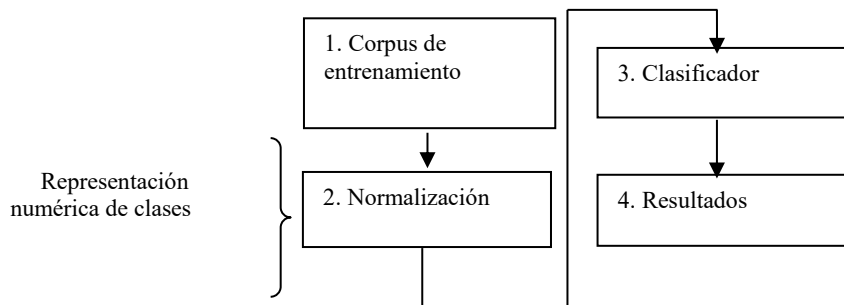


Fig. 1. Metodología propuesta.

4. Experimentación

La primera aproximación de solución consistió en utilizar el corpus de entrenamiento importando a Excel el archivo en formato XML. Se eliminó la información que no se requiere para el análisis de sentimientos, permaneciendo solamente la columna etiquetada con el texto *content*, que es el tuit a analizar y la columna *value*, que contiene las etiquetas de las clases (+P, P, NEU, N, N+ y NONE).

Con este formato se utilizaron los clasificadores J48 y Naive Bayes, posteriormente se cambió el formato de la columna *value* para que las clases utilizaran una representación numérica de la siguiente forma: 0 = NONE, 1 = NEU, 2 = N, 3 = N+, 4 = P y 5 = +P y poder experimentar con el clasificador MSV. Se inicia el programa de análisis de datos Weka versión 3.6 y se importó el archivo .csv que contiene los datos de entrenamiento etiquetados del corpus TASS 2015 otorgándonos una vista de la distribución de clases. Se seleccionan los clasificadores J48, LibSVM y NaiveBayes con el parámetro porcentaje de división al 80% (*percentage split*), que se refiere a que, de los datos ingresados, se tomará el 80% de entrenamiento y el 20% restante como datos de prueba. Los resultados de este experimento se muestran en la tabla 2:

Tabla 2. Resultados de clasificación en seis categorías, sin pre procesamiento, utilizando 80% de división.

Clasificador	J48	LibSVM	Naive Bayes
Resultado	.2458	.2437	.2458

El segundo experimento consistió en utilizar la validación cruzada de 10 folders eligiendo los mismos clasificadores sin pre-procesamiento. Los resultados se muestran en la tabla 3.

Tabla 3. Resultados de clasificación en seis categorías, sin pre procesamiento, utilizando validación cruzada de 10 folders.

Clasificador	J48	LibSVM	Naive Bayes
Resultado	.2317	.2288	.2317

Para la construcción del segundo modelo, se dividió el corpus de entrenamiento en 6 diferentes archivos, cada uno, contiene los textos que corresponden a una clase. Se programaron componentes en Python utilizando expresiones regulares para extraer características estadísticas y léxicas, como la frecuencia por tuit de hashtags, urls, y menciones a usuario; de las características léxicas, se utilizó una lista de emoticones de Wikipedia, categorizados como Sonrisa, Risa y Tristeza [17], con el fin de obtener una frecuencia por tuit. El estado del arte indica que el uso de mayúsculas representa énfasis al tratar de transmitir una idea, opinión o sentimiento, de tal forma que se tomaron en cuenta las frecuencias por cada tuit de las palabras que comienzan en mayúscula, así como palabras que se encuentran escritas en su totalidad en mayúsculas. El archivo de características es utilizado para ingresarlo a Weka, construir el modelo de clasificación con los mismos parámetros que el modelo descrito anteriormente; con el fin de comparar los resultados obtenidos. Cabe hacer mención, que los tuits hasta este momento, no han recibido ningún tipo de pre procesamiento.

Tabla 4. Resultados de clasificación en seis categorías, sin pre procesamiento, utilizando 80% de división con las características léxicas extraídas.

Clasificador	J48	LibSVM	Naive Bayes
Resultado	.2832	.2853	.2610

Utilizando las mismas características, pero evaluando mediante validación cruzada de 10 folders, se obtienen los siguientes resultados:

Tabla 5. Resultados de clasificación en seis categorías, sin pre procesamiento, utilizando validación cruzada de 10 folders con las características léxicas extraídas.

Clasificador	J48	LibSVM	Naive Bayes
Resultado	.2811	.2878	.2577

5. Comparación de resultados

En la figura 2, se muestran los resultados obtenidos en este trabajo y se comparan con los alcanzados por los equipos participantes en el TASS 2015, que obtuvieron los cuatro primeros lugares en el taller y los obtenidos en nuestros experimentos. Los

experimentos son evaluados con la métrica de exactitud (Accuracy), de acuerdo a lo estipulado en la página oficial del concurso TASS 2015¹.

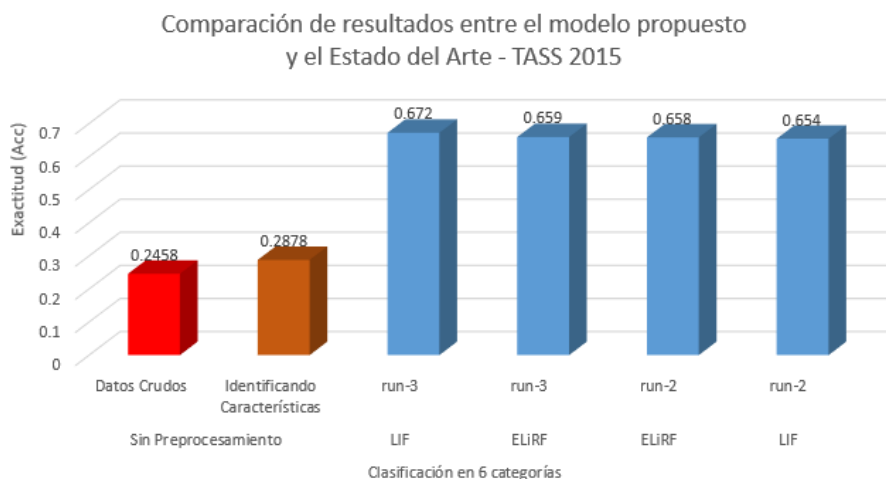


Fig. 2. Comparación de resultados con los del estado del arte.

Como se puede observar, el modelo propuesto aún necesita ser ajustado para alcanzar mejores resultados.

6. Conclusiones y trabajo futuro

El estado del arte indica que las características especiales del lenguaje de Twitter, requieren un tratamiento especial para analizar los textos. La sintaxis particular, menciones a usuarios, urls, hashtags, emoticones, oraciones pobres en gramática, modismos, entre otras, conducen a una caída en el rendimiento de las herramientas tradicionales del PLN [18]. Existe una propuesta en los trabajos [19] y [20] para la normalización del texto que presentan estas características.

Este artículo resume la experimentación realizada con los datos del corpus TASS 2015, utilizando como primera aproximación los clasificadores J48, Naive Bayes y MSV incluidos en Weka. Se observó que, al extraer frecuencias de características, los resultados han mejorado. El estado del arte indica que es importante el pre procesamiento en este tipo de textos, por lo que en primera instancia podemos concluir, que la etapa de pre procesamiento puede ser un factor clave en el tratamiento de la información para alcanzar mejores resultados.

Como trabajo futuro, se propone agregar como características las frecuencias de término y la frecuencia inversa del término (*tf-idf*), representación mediante unigramas y bigramas experimentando con los mismos clasificadores y los mismos parámetros, comparando el rendimiento de los modelos en base a sus resultados. Se considera

¹ <http://www.sepln.org/workshops/tass/2015/tass2015.php>

importante, que los experimentos futuros incluyan algún tipo de pre procesamiento, por lo tanto se pretende utilizar la herramienta desarrollada por [23], con el fin de apoyar esta etapa buscando más características léxicas para mejorar los resultados hasta ahora alcanzados.

Agradecimientos. Trabajo soportado por CONACYT.

Referencias

1. Cesteros, J. S., Almeida, A., de Ipiña, D. L.: DeustoTech Internet at TASS 2015: Sentiment Analysis and Polarity Classification in Spanish Tweets. In: TASS@ SEPLN, 2, 23–28 (2015)
2. Rodríguez, M., Alvarez, J., Mesa, J., González, A.: Metodologías para la realización de proyectos de Data Mining. Ponencia presentada en el VII Congreso Internacional de Ingeniería de ProyectosK Elissa (2005)
3. Valverde, J., Tejada, J., Cuadros, E.: Comparing Supervised Learning Methods for Classifying Spanish Tweets. CEUR-WS, 87–92 (2015)
4. Le, Quoc V., Mikolov, T.: Distributed representations of sentences and documents. In: ICML, 14, 1188–1196 (2014)
5. Zafra, J., María, S., Martínez Cámara, E., Valdivia, M., Teresa, M., Ureña López, L. A.: SINAI-ESMA: An unsupervised approach for Sentiment Analysis in Twitter. In: Proceedings of the TASS workshop at SEPLN, 16–19 (2014)
6. Hernández Petlachi, R., Xiaou, L.: Análisis de sentimiento sobre textos en español basado en aproximaciones semánticas con reglas lingüísticas. In: Proceedings of the TASS workshop at SEPLN (2014)
7. Villena-Román, J., Garcia-Morera, J., Garcia-Cumbreras, M. A., Martinez-Cámara, E., Martín-Valdivia, M. T., Ureña-López, L. A.: Overview of TASS 2015. In: TASS 2015 Workshop on Sentiment Analysis at SEPLN, 1397, 13–21 (2015)
8. Liu, B.: Sentiment Analysis, Mining opinions, sentiments and emotions. Cambridge University Press, First edition (2015)
9. Rouvier, M., Benoit, F.: LIF @ TASS 2015: Deep models are very cooooooolllll for sentiment analysis (2015)
10. Hurtado, L. F., Pla, F., Buscaldi, D.: ELiRF-UPV en TASS 2015: Análisis de Sentimientos en Twitter. In: Proceedings of TASS 2015, Workshop on Sentiment Analysis at SEPLN, 35–40 (2015)
11. Sepln.org.: TASS 2016. @ SEPLN, Recuperado de: <http://www.sepln.org/workshops/tass/2016/tass2016.php> Fecha de Acceso 30 Agosto de 2016 (2016)
12. Saralegi, X., San Vicente, I.: Tass: Detecting sentiments in Spanish tweets. In: Workshop on Sentiment Analysis at SEPLN (TASS), SEPLN, 9 (2012)
13. Fernández, J., Gutiérrez, Y., Gómez, J. M., Martínez-Barco, P., Montoyo, A., Muñoz, R.: Sentiment analysis of Spanish tweets using a ranking algorithm and skipgrams (2013)
14. Urizar, X. S., Roncal, I. S. V.: Elhuyar at TASS 2013. In: Proceedings of the Workshop on Sentiment Analysis at SEPLN (TASS 2013), 143–150 (2013)
15. San Vicente Roncal, I., Urizar, X. S., Looking for features for supervised tweet polarity classification. In Proceedings of the TASS workshop at SEPLN (2014)
16. Liu, B.: Web data mining: exploring hyperlinks, contents, and usage data. Springer Science & Business Media (2007)
17. Anexo: Emoticonos (s.f.). En Wikipedia, Recuperado el 16 de septiembre de 2016 de <https://es.wikipedia.org/wiki/Anexo:Emoticonos>
18. San Vicente Roncal, I., Urizar, X. S.: Looking for features for supervised tweet polarity classification. In: Proceedings of the TASS workshop at SEPLN (2014)

19. Brody, S., Diakopoulos, N.: Cooooooooooooolllllllllllll!!!!!!!!!!!!!: using word lengthening to detect sentiment in microblogs. In Proceedings of the conference on empirical methods in natural language processing, 562–570 (2011)
20. Han, B., Baldwin, T.: Lexical normalization of short text messages, Makn sens a# twitter. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies 1, 368–378 (2011)
21. Sanzón, Y. M., Vilariño, D., Somodevilla, M. J., Zepeda, C., Tovar, M.: Modelos para detectar la polaridad de los mensajes en redes sociales. Research in Computing Science, Avances en la Ingeniería del Lenguaje y del Conocimiento, 29 (2015)
22. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Foundations and trends in information retrieval, 2(1-2), 1–135 (2008)
23. Sidorov, G., Miranda-Jiménez, S., Viveros-Jiménez, F., Gelbukh, A., Castro-Sánchez, N., Velásquez, F., Gordon, J.: Empirical study of machine learning based approach for opinion mining in tweets. In: Mexican International Conference on Artificial Intelligence, 1–14 Springer Berlin Heidelberg (2012)
24. Liu, B.: Sentiment Analysis and Subjectivity. Handbook of natural language processing, 2, 627–666 (2010)