Procesamiento del Lenguaje Natural, Revista nº 51, septiembre de 2013, pp 127-134 recibido 02-04-2013 revisado 16-06-2013 aceptado 21-06-2013

Una aproximación supervisada para la minería de opiniones sobre tuits en español en base a conocimiento lingüístico

A supervised approach to opinion mining on Spanish tweets based on linguistic knowledge

David Vilares, Miguel A. Alonso y **Carlos Gómez-Rodríguez**
Departamento de Computación, Universidade da Coruña
Campus de Elviña, 15011 A Coruña
{david.vilares, miguel.alonso, carlos.gomez}@udc.es

Resumen: En este artículo se describe un sistema para la clasificación de la polaridad de tuits escritos en español. Se adopta una aproximación híbrida, que combina conocimiento lingüístico obtenido mediante PLN con técnicas de aprendizaje automático. Como paso previo, se realiza una primera etapa de preprocesado para tratar ciertas características del uso del lenguaje en Twitter. A continuación se extrae información morfológica, sintáctica y semántica, para utilizarla posteriormente como entrada a un clasificador supervisado. La evaluación de la propuesta se lleva a cabo sobre el corpus TASS 2012, anotado para realizar tareas de clasificación con cuatro y seis categorías. Los resultados experimentales muestran un buen rendimiento para ambos casos, lo que valida la utilidad práctica de la propuesta.
Palabras clave: Análisis del sentimiento, Minería de opiniones, Análisis sintáctico de dependencias, Aprendizaje automático, Twitter

Abstract: This article describes a system that classifies the polarity of Spanish tweets. We adopt a hybrid approach, which combines linguistic knowledge acquired by means of NLP with machine learning techniques. We carry out a preprocessing of the tweets as an initial step to address some characteristics of the language used in Twitter. Then, we apply part-of-speech tagging, dependency parsing and extraction of semantic knowledge, and we employ all that information as features for a supervised classifier. We have evaluated our proposal with the TASS 2012 corpus, which is annotated to undertake classification tasks with four and six categories. Experimental results are good in both cases and confirm the practical utility of the approach.
Keywords: Sentiment Analysis, Opinion Mining, Dependency Parsing, Machine Learning, Twitter

1. Introducción

Con la explosión de la Web 2.0 muchos usuarios emplean los medios sociales para compartir sus opiniones y experiencias acerca de productos, servicios o personas. Esta situación ha despertado un gran interés a nivel empresarial, ya que se ve en estos recursos un mecanismo para conocer, de manera eficaz y global, el punto de vista de consumidores sobre una gran variedad de temas. Sin embargo, el análisis manual de este tipo de medios no es una solución viable, dado el flujo ingente de opiniones que en ellos se expresan. A este respecto, la minería de opiniones (MO), conocida también como análisis del sentimiento; es una reciente área de investigación centrada en tareas como determinar automáticamente si en un texto se opina o no, o si la *polaridad* o *sentimiento* que se expresa en él es positiva, negativa o mixta. También es útil de cara a la extracción automática de características, lo que permite conocer la percepción que se tiene sobre aspectos concretos de un tema (p. ej. *"La película Y tiene un gran final"*).

En este contexto, una de los medios sociales más populares es Twitter. En esta red de microblogging, los usuarios expresan sus opiniones en mensajes de hasta 140 caracteres, especialmente sobre temas de actualidad, lo que constituye una importante fuente de información desde el punto de vista de la inteligencia de negocio.

En este artículo presentamos una aproximación que combina conocimiento morfológico, sintáctico y semántico con técnicas de

ISSN 1135-5948 © 2013 Sociedad Española Para el Procesamiento del Lenguaje Natural

aprendizaje automático para tratar de clasificar la opinión de tuits escritos en español. Para evaluar nuestra propuesta se ha empleado el corpus TASS 2012, donde se distinguen hasta seis categorías distintas.

El resto del artículo se organiza como sigue. En la sección 2 se revisa brevemente el estado del arte en lo referido a la clasificación de la polaridad, centrándonos en estudios relacionados con Twitter. En la sección 3 detallamos nuestra aproximación. En la sección 4 se ilustran los resultados experimentales. Por último, en la sección 5 se presentan las conclusiones y las líneas de trabajo futuras.

2. Estado del arte

Desde que la MO ha sido asumida como un nuevo reto por parte del área de PLN, una de las tareas en las que más esfuerzos se han depositado es la clasificación de la polaridad. Dos han sido los principales enfoques propuestos para resolver este problema: el semántico y el supervisado. El primero (Turney, 2002), se apoya en diccionarios semánticos, conocidos también como lexicones de opiniones, donde a cada palabra que denota opinión se le asigna una orientación semántica (OS). El segundo enfoque (Pang, Lee, y Vaithyanathan, 2002), propone una solución basada en aprendizaje automático (AA), asumiendo la tarea como un problema genérico de clasificación.

La MO se ha centrado tradicionalmente en el análisis de textos extensos. Sin embargo, en los últimos años se ha prestado especial atención a la detección del sentimiento en microtextos, sobre todo en lengua inglesa, dado el éxito de las redes de microblogging. Por ejemplo, Thelwall et al. (2010) proponen el algoritmo SentiStrength para el análisis de textos cortos en la red social MySpace. En (Thelwall, Buckley, y Paltoglou, 2011) se estudia cómo los eventos sociales influyen en el número de opiniones expresadas en Twitter. En (Thelwall, Buckley, y Paltoglou, 2012) se extiende y adapta SentiStrength a otros medios de comunicación web de mensajes cortos. Por otro lado, Bakliwal et al. (2012) presentan un método de puntuación del sentimiento no supervisado y comparan su propuesta con una aproximación supervisada, alcanzando una precisión similar. Zhang et al. (2011) proponen un método supervisado para el análisis de tuits que se entrena a partir de los datos proporcionados por un analizador semántico base no supervisado.

Un problema del análisis del sentimiento sobre textos cortos es el coste que presenta la creación de corpora y de recursos específicos. Para el castellano, destaca el corpus TASS 2012, presentado en el Taller de Análisis del Sentimiento en la SEPLN (Villena-Román et al., 2013), y que utilizamos en este artículo. Se trata de una colección de tuits en español escritos por diversas personalidades públicas. Está constituido por un conjunto de entrenamiento y uno de test que contienen 7,219 y 60,798 tuits respectivamente. Cada uno de ellos está anotado con una de la siguientes categorías: *muy positivo* (P+), *positivo* (P), *neutro/mixto* (NEU), *negativo* (N), *muy negativo* (N+) o *sin opinión* (NONE). En relación con la detección del sentimiento, el taller propuso dos actividades: una de clasificación en seis categorías (con las seis clases mencionadas) y otra de clasificación en cuatro (las clases P+ y N+ se incluyen en las clases P y N respectivamente). La etiquetación del conjunto de test se caracteriza por haber sido obtenida mediante un *pooling* de los resultados de los sistemas que participaron en el taller; seguida de una revisión manual para los casos conflictivos.

Existen diversas propuestas que han evaluado este corpus, la mayoría de ellas enmarcadas dentro de un enfoque basado en AA. Por ejemplo, Saralegi Urizar y San Vicente Roncal (2012) plantean una solución supervisada que emplea conocimiento lingüístico para obtener los atributos de entrada al clasificador. Realizan tareas de lematización y etiquetación, y consideran aspectos relevantes en un entorno de MO, como son los emoticonos o la negación. Batista y Ribeiro (2013) plantean una aproximación supervisada. Sin embargo, en lugar de entrenar un clasificador que distinga n polaridades, construyen *n-1* clasificadores binarios. Finalmente combinan los resultados de los modelos entrenados para eliminar ambigüedades en la clasificación de ciertos tuits. Trilla y Alías (2012) proponen un esquema de clasificación de texto basado en el Multinomial Naive Bayes. Martínez Cámara et al. (2012) también plantean una solución de AA supervisado, en concreto aplicando Support Vector Machines (SVM). Como entrada al clasificador usan una bolsa de unigramas que representa el mensaje. Además incluyen otras características como emotico-

nos, o el número de palabras positivas y negativas presentes en un tuit. En la misma líneas, Fernández Anta et al. (2012) comparan el rendimiento de varios clasificadores supervisados sobre el corpus TASS 2012. Por otra parte, Martín-Wanton y Carrillo de Albornoz (2012) proponen un método basado diccionarios afectivos y WordNet, mientras que Castellano González, Cigarrán Recuero, y García Serrano (2012) abordan el problema del análisis del sentimiento desde una perspectiva de recuperación de información. Por último, Moreno-Ortiz y Pérez-Hernández (2012) emplean Sentitext, un sistema léxico no supervisado para el análisis del sentimiento de textos en castellano, para clasificar la polaridad.

3. *Sistema híbrido de clasificación de polaridades*

En este artículo se propone una aproximación que combina conocimiento lingüístico con técnicas supervisadas para tratar de resolver los problemas que ambos enfoques presentan separadamente, y que se ven acentuados en un medio como Twitter. Las aproximaciones semánticas se caracterizan por emplear un diccionario de OS genérico. Este enfoque ha demostrado ser útil en distintos ámbitos, pero su rendimiento en Twitter disminuye ya que en este medio existe una elevada frecuencia de abreviaturas, emoticonos o expresiones que denotan opinión, pero cuyas OS no se encuentran en un lexicón genérico, lo que en términos de rendimiento se traduce en un bajo *recall* (Zhang et al., 2011). Respecto a las aproximaciones supervisadas, el principal problema reside en su dependencia del dominio y en el coste de crear conjuntos de entrenamiento. Estos métodos representan el texto como una bolsa de palabras, aprendiendo satisfactoriamente la percepción de un término para un ámbito en concreto. Sin embargo, su rendimiento cae drásticamente al clasificar textos de un campo distinto (Taboada et al., 2011).

Nuestra propuesta parte de un sistema base, para a continuación determinar cómo la información morfosintáctica sirve de ayuda a un clasificador supervisado. Además, se propone un método automático para maximizar el rendimiento en un dominio en concreto, atendiendo a criterios semánticos. Como clasificador, se ha optado por una SMO, una implementación de SVM presentada en (Platt, 1999), y que se incorpora en el software de algoritmos de aprendizaje automático WEKA (Hall et al., 2009). La elección de este clasificador se debe al buen rendimiento que se obtuvo con él sobre el corpus de entrenamiento y a los buenos resultados que otros estudios han conseguido sobre el mismo corpus (Saralegi Urizar y San Vicente Roncal, 2012).

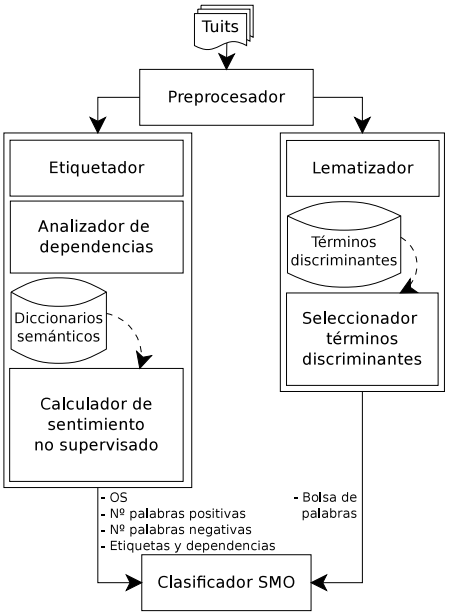

Figura 1: Arquitectura general del sistema

La figura 1 describe la arquitectura general del sistema, cuyo funcionamiento pasamos a describir en los siguientes subapartados.

3.1. Preprocesado

Como paso previo, todos los tuits fueron sometidos al siguiente preprocesado *ad-hoc*, para tratar el uso particular que se hace del lenguaje en Twitter:

- *Tratamiento de emoticonos*: Existe una gran variedad de símbolos que se emplean para reflejar un estado de ánimo. Para identificarlos se utilizó la colección recogida en (Agarwal et al., 2011). Cada emoticono se sustituye por una de estas cinco etiquetas: muy positivo (EMP), positivo (EP), neutro (ENEU), negativo (EN) y muy negativo (EMP).
- *Normalización de URL's*: Las direcciones web presentes en un tuit son sustituidas por la cadena *"URL"*.
- *Corrección de abreviaturas más frecuentes*: Se sustituyen algunos de los vocablos no gramaticales más habituales (p.

ej. *"q"*, *"xq"*,...) por su forma reconocida.

- *Normalización de risas*: Las expresiones típicas que permiten reflejar este fenómeno vía escrita (p. ej. *"jajjja"*, *"JJJEJEJE"*,...), son normalizadas como *jxjx* donde $x \in \{a, e, i, o, u\}$.

- *Tratamiento de elementos específicos de Twitter ("@"y "#")*: Las menciones al usuario se mantienen eliminando la *"@"* y capitalizando la primera letra (p. ej. *"@usuario"* pasa a ser *"Usuario"*).[1] Respecto a los hashtags (p. ej. *"#sepln"*), si aparece al principio o al final del tuit se elimina el mismo. En caso contrario se suprime solamente la *"#"* (p. ej. *"#sepln"* pasa a ser *"sepln"*).

3.2. Propuesta base

Como sistema base se utilizó una aproximación semántica presentada en (Vilares, Alonso, y Gómez-Rodríguez, 2013). Se realizan tareas de segmentación, tokenización y etiquetación morfológica para luego obtener el *árbol de dependencias* de cada oración mediante algoritmos de análisis sintáctico de dependencias, utilizando para ello MaltParser (Nivre et al., 2007). Este tipo de análisis establece vínculos padre/dependiente entre pares de palabras. A cada uno de esos vínculos se les denomina *dependencia* y son anotados con la función sintáctica que relaciona los dos términos. En la figura 2 se ilustra un ejemplo de este tipo de análisis.

El árbol de dependencias se emplea entonces para realizar el análisis del sentimiento sobre los tuits, tratando sintácticamente tres de las construcciones más significativas en el ámbito de la MO: la intensificación, las oraciones subordinadas adversativas y la negación. Para determinar la OS de las palabras que denotan opinión, se utiliza el SO-DictionariesV1.11Spa, una colección de diccionarios semánticos genéricos presentados en (Brooke, Tofiloski, y Taboada, 2009). Como resultado del análisis semántico se obtiene la OS global del texto. Este valor, junto con el número total de palabras positivas y negativas presentes en un tuit, se emplearán como atributos para entrenar un clasificador.

[1]La letra mayúscula se fuerza para que se trate como un nombre propio en la etapa de análisis morfológico.

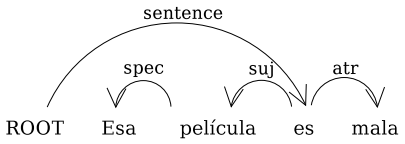

Figura 2: Ejemplo de un árbol de dependencias.

3.3. Información morfosintáctica (IMS)

La utilización de etiquetas morfológicas como un elemento que por sí mismo sirva de ayuda en la clasificación de la polaridad de un texto es una cuestión que ha sido discutida en varias ocasiones. Pak y Paroubek (2010) observaron que la distribución de etiquetas en textos de microblogging es distinta según se trate de un mensaje positivo, negativo o informativo. En la misma línea, Spencer y Uchyigit (2012) y Saralegi Urizar y San Vicente Roncal (2012) destacan una mayor presencia de ciertas categorías gramaticales en textos subjetivos respecto a los objetivos, como es el caso de los adjetivos o las interjecciones. En el corpus TASS 2012 también se observó que algunas etiquetas son más habituales dependiendo de la categoría del tuit. En la tabla 1 mostramos la frecuencia de aparición para algunas de ellas. Para tratar este fenómeno, el número total de apariciones de cada etiqueta en un tuit se incorporará como atributo de entrada a nuestro clasificador.

Cat.	a	n	v	i	f
P+	0,060	0,256	0,111	0,004	0,215
P	0,056	0,266	0,119	0,002	0,198
NEU	0,057	0,254	0,133	0,001	0,163
N	0,050	0,263	0,132	0,001	0,161
N+	0,060	0,266	0,118	0,001	0,154
NONE	0,048	0,299	0,090	0,001	0,220

Tabla 1: Frecuencia de etiquetas en el conjunto de entrenamiento: adjetivos (a), nombres (n), verbos (v), interjecciones (i) y signos de puntuación (f)

Del mismo modo, creemos que ciertos tipos de funciones sintácticas pueden ser más habituales en un tuit en función de su polaridad. En la tabla 2 se muestra la frecuencia de algunas dependencias en los tuits del corpus TASS 2012. Destaca el empleo de dependencias como la de complemento agente (cag), que sugiere que, en las opiniones negativas, la forma pasiva es más habitual. Siguiendo

la misma estrategia que para el caso de las etiquetas, el número de apariciones de cada tipo de dependencia en un tuit se incluirá como atributos de entrada al clasificador.

Cat.	ci	atr	cc	cag
P+	0,008	0,105	0,042	0,004
P	0,010	0,010	0,051	0,000
NEU	0,010	0,141	0,053	0,001
N	0,009	0,000	0,055	0,150
N+	0,007	0,000	0,049	0,145
NONE	0,179	0,008	0,003	0,001

Tabla 2: Frecuencia de etiquetas sintácticas en el conjunto de entrenamiento: complemento indirecto (*ci*), atributo (*atr*), complemento circunstancial (*cc*) y complemento agente (*cag*)

3.4. Adaptación al dominio (AD)

La utilización de conocimiento semántico genérico permite obtener un buen rendimiento base. Sin embargo, estos lexicones presentan varios problemas. Uno de los inconvenientes es su baja cobertura, limitada a los términos incluidos en ellos. Otro es el referido a las OS asignadas a cada término. Éstas, al adaptarse a un contexto general, pueden ser incorrectas para ciertos dominios y contextos, lo que repercute negativamente en el rendimiento. Por ejemplo, la palabra *"asesino"* se percibe intuitivamente como negativa, pero dentro de un dominio de películas, probablemente ese término sirva para describir la temática de la misma, pero no su calidad.

También deben considerarse las peculiaridades del medio en el que se se opina. Los mensajes publicados en Twitter incluyen, con frecuencia, elementos que expresan subjetividad u opinión, pero cuya OS no se refleja en un lexicón de opiniones. Un ejemplo es el uso de emoticonos (*":)"*, *":("*,...) o de interjecciones (*"jaja"*, *"jeje"*,...). Del mismo modo, la utilización de elementos como la etiqueta *Follow Friday* (*"FF"*) o la difusión del mensaje de otro usuario (*"RT"*), llevan asociada con frecuencia una carga subjetiva.

Nuestra solución consiste en desarrollar un mecanismo que permita enriquecer y adaptar automáticamente el conocimiento semántico a las características del dominio. Para ello, fueron dos las tareas que se abordaron: *selección de los términos más discriminantes* y *modificación de los diccionarios genéricos*.

3.4.1. Selección de los términos más discriminantes

El objetivo es crear una bolsa de palabras formada por términos que permitan distinguir bien entre las distintas categorías; de forma que cada uno de esos vocablos se incorporen como entradas del clasificador. Sobre el esquema de pesado utilizado para estos atributos, se ha optado por una ocurrencia binaria, dado que es poco habitual que una misma palabra se repita dentro de un tuit.

Para conocer los elementos que pueden constituir dicha bolsa de palabras, se ha utilizado el soporte de selección de atributos de WEKA y el conjunto de entrenamiento del corpus TASS 2012. Como paso previo se preprocesaron y lematizaron los tuits. Los términos resultantes se han clasificado ordenadamente en función de la ganancia de información que proporcionan respecto a la clase. Para que dicha selección fuese más robusta, se llevó a cabo una validación cruzada de 10 iteraciones. La lista de vocablos resultantes, para las actividades de clasificación en cuatro y seis clases, está constituida por más de 14.000 elementos. Sin embargo, se ha comprobado empíricamente que las mejores configuraciones, en términos de rendimiento, se obtienen con un número reducido de palabras (entre 100 y 400). En la tabla 3 se muestran algunos elementos característicos de esta red de microblogging y la posición que ocupan en la clasificación.

Término	Posición (4 clases)	Posición (6 clases)
EP (emoticono)	1	1
URL	4	4
FF	30	47
jaja	101	11.964
EN (emoticono)	70	649

Tabla 3: Relevancia de términos para la detección del sentimiento en el corpus TASS 2012

3.4.2. Modificación de los diccionarios genéricos

Como en el apartado anterior, se han extraído un listado de los elementos más discriminantes, del corpus de entrenamiento, aunque en este caso solo se consideraron los tuits con polaridad positiva (P, P+) o negativa (N, N+). De esta manera se conocen los elementos polares más representativos. A continuación se

asigna, automáticamente, una OS adaptada (OSA) a cada uno de ellos; en función de sus apariciones en textos positivos y negativos, y se aplica el siguiente conjunto de reglas:

1. Si la OSA es inferior a 0,5 en valor absoluto entonces el término correspondiente no modifica el lexicón de opiniones.
2. Si el vocablo ya existe en los diccionarios semánticos genéricos y sus respectivas OSA y OS genérica tienen el mismo signo de polaridad, entonces el valor de esta última prevalece.
3. Si las dos reglas anteriores no se cumplen, el término correspondiente pasa a formar parte del diccionario con su OSA.

Con este método se extrajeron más de 10.000 nuevos términos polares, que fueron ordenados, como en el caso anterior, en función de la ganancia de información que proporcionan respecto a la clase. Además, los diccionarios semánticos también se adaptaron para resolver el problema del principio de cortesía (Brown y Levinson, 1987). Existe una tendencia humana a suavizar las críticas negativas mediante la utilización de eufemismos o negaciones de términos positivos. Para compensarlo, algunos sistemas léxicos (Taboada et al., 2011) potencian la OS de los elementos negativos. En nuestra propuesta también se determinó, empíricamente, que incrementar en un 100 % la OS de este tipo de términos, mejora ligeramente el rendimiento.

El enriquecimiento y adaptación de los diccionarios semánticos genéricos logró un incremento notable cuando fue considerado aisladamente. Sin embargo, su efecto se disipó casi por completo al aplicarlo conjuntamente junto con la selección de los términos más relevantes. De todos modos, se comprobó empíricamente que, incluir entre el 50 % y el 70 % del total de los términos polares negativos extraídos, tiene un ligero beneficio en el rendimiento; lo que refuerza la hipótesis de la tendencia positiva en el lenguaje humano.

4. Resultados experimentales

Para la evaluación de nuestra propuesta hemos utilizado el conjunto de test del corpus TASS 2012. Se han realizado experimentos para las dos tareas de clasificación de polaridad propuestas en el taller: cuatro categorías (P, NEU, N y NONE) y seis categorías (P+, P, NEU, N, N+ y NONE). La tablas 4 y 5 ilustran los resultados para ambas tareas, desglosados según las distintas versiones desarrolladas. En ambos casos, la propuesta inicial consigue un buen rendimiento. La incorporación de la información morfosintáctica mejora la detección de la polaridad para textos positivos y negativos. Ello refuerza la idea de que los usuarios tienden a emplear ciertas etiquetas y patrones sintácticos según el tipo de opinión a transmitir. La *accuracy* obtenida en la versión final sugiere que, aunque con los lexicones de opiniones genéricos y la morfosintaxis se alcanza un buen rendimiento, es necesario incorporar conocimiento semántico especifico del dominio para optimizar la propuesta.

Medida	**Base**	**+ims**	**+ad**
F_{p}	0,631	0,680	0,745
F_{neu}	0,000	0,000	0,054
F_{n}	0,566	0,603	0,671
F_{none}	0,574	0,564	0,620
Accuracy	0,587	0,615	0,676

Tabla 4: Resultados obtenidos sobre la evaluación del conjunto de test (4 clases)

Medida	**Base**	**+ims**	**+ad**
F_{p+}	0,609	0,637	0,705
F_{p}	0,000	0,040	0,307
F_{neu}	0,000	0,009	0,089
F_{n}	0,452	0,478	0,512
F_{n+}	0,000	0,120	0,441
F_{none}	0,575	0,605	0,648
Accuracy	0,523	0,546	0,600

Tabla 5: Resultados obtenidos sobre la evaluación del conjunto de test (6 clases)

En ambas tareas, la clasificación de los tuits neutros alcanza un bajo rendimiento, lo que creemos que es debido a dos factores. El primero está relacionado con una característica intrínseca de este tipo de críticas: la mezcla de ideas a favor y en contra dificulta la clasificación de estos textos, que pueden ser confundidos con opiniones positivas o negativas; más aún cuando se trata de tuits, donde un usuario no dispone de espacio para desarrollar su argumento. El segundo es referido al criterio de clasificación establecido en el corpus TASS 2012, donde el límite entre un tuit neutro y uno sin opinión, o con polaridad positiva o negativa, es difuso. Ello

afecta negativamente a las tareas de clasificación, dada la falta de un criterio objetivo que permita diferenciar las distintas polaridades. En una línea similar, este problema ya ha sido comentado por otros autores que han trabajado sobre el mismo corpus (Saralegi Urizar y San Vicente Roncal, 2012). La tabla 6 compara la *accuracy* de nuestra aproximación con las de los participantes del TASS 2012. Algunos de ellos enviaron varias propuestas al taller, aunque aquí solo se indica aquella con la que obtuvieron un mayor rendimiento. En (Villena-Román et al., 2013) se encuentran en detalle los resultados para cada una de ellas.

Propuesta	***Acc.* 4 cat**	***Acc.* 6 cat**
ELHUYAR	0,711	0,653
L2F-INESC	0,691	0,622
Nuestra propuesta	0,676	0,600
LA SALLE-URL	0,619	0,570
SINAI-UJAEN	0,606	0,549
LSI UNED	0,590	0,538
LSI UNED2	0,501	0,404
IMDEA	0,459	0,360
UMA	0,351	0,167

Tabla 6: Comparativa con los participantes del TASS 2012

5. Conclusiones y trabajo futuro

Este artículo presenta una propuesta que emplea conocimiento lingüístico para entrenar un clasificador que detecte el sentimiento de tuits escritos en español. Los resultados experimentales muestran un buen rendimiento y sugieren que la estructura morfosintáctica de los textos es útil para detectar la polaridad.

De cara al futuro hay varios aspectos que nos gustaría explorar. El preprocesamiento actual tuits es bastante simple. Nos gustaría determinar cómo una normalización gramatical de los tuits podría ayudar a clasificar la polaridad. A este respecto, en (Oliva et al., 2013) se propone un sistema de normalización de SMS que podría servir para enriquecer nuestro preprocesado. También creemos que puede ser de utilidad integrar en nuestra aproximación la propuesta de (Batista y Ribeiro, 2013), donde se propone entrenar varios clasificadores binarios y combinar los resultados obtenidos, explotando así las diferencias entre textos que expresan sentimientos distintos. La forma de emplear las dependencias sintácticas como entrada al clasificador también es un aspecto en el que nos gustaría profundizar. Actualmente solamente se utiliza como atributos el número total de cada tipo de dependencia en un tuit. Nos gustaría explorar cómo utilizar las tripletas padre/dependencia/dependiente puede ayudar en tareas de clasificación de polaridad.

Agradecimientos

Trabajo parcialmente financiado por el Ministerio de Economía y Competitividad y FEDER (TIN2010-18552-C03-02) y por la Xunta de Galicia (CN2012/008, CN 2012/319).

Bibliografía

Agarwal, A., B. Xie, I. Vovsha, O. Rambow, y R. Passonneau. 2011. Sentiment analysis of Twitter data. En *Proceedings of the Workshop on Languages in Social Media*, LSM '11, páginas 30–38, Stroudsburg, PA, USA. ACL.

Bakliwal, A., P. Arora, S. Madhappan, N. Kapre, M. Singh, y V. Varma. 2012. Mining sentiments from tweets. En *Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis*, páginas 11–18, Jeju, Korea. ACL.

Batista, F. y R. Ribeiro. 2013. The L2F Strategy for Sentiment Analysis and Topic Classification. *Procesamiento de Lenguaje Natural*, 50:77–84.

Brooke, J., M. Tofiloski, y M. Taboada. 2009. Cross-Linguistic Sentiment Analysis: From English to Spanish. En *Proceedings of the International Conference RANLP-2009*, páginas 50–54, Borovets, Bulgaria. ACL.

Brown, P. y S Levinson. 1987. *Politeness, Some universals in language use*. Cambridge, Cambridge University Press.

Castellano González, A., J. Cigarrán Recuero, y A. García Serrano. 2012. Using IR techniques for topic-based sentiment analysis through divergence models. En *TASS 2012 Working Notes*, Castellón, Spain.

Fernández Anta, A., P. Morere, L. Núñez Chiroque, y A. Santos. 2012. Techniques for Sentiment Analysis and Topic Detection of Spanish Tweets: Preliminary Report. En *TASS 2012 Working Notes*, Castellón, Spain.

Hall, M., E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, y I.H. Witten. 2009. The WEKA data mining software: an update. *SIGKDD Explor. Newsl.*, 11(1):10–18, Noviembre.

Martín-Wanton, T. y J. Carrillo de Albornoz. 2012. Sistema para la Clasificación de la Polaridad y Seguimiento de Temas. En *TASS 2012 Working Notes*, Castellón, Spain.

Martínez Cámara, E., M. T. Martín Valdivia, M. A. García Cumbreras, y L. A. Ureña López. 2012. SINAI at TASS 2012. *Procesamiento de Lenguaje Natural*, 50:53–60.

Moreno-Ortiz, A. y C. Pérez-Hernández. 2012. Lexicon-Based Sentiment Analysis of Twitter Messages in Spanish. En *TASS 2012 Working Notes*, Castellón, Spain.

Nivre, Joakim, Johan Hall, Jens Nilsson, Atanas Chanev, Gülsen Eryigit, Sandra Kübler, Svetoslav Marinov, y Erwin Marsi. 2007. Maltparser: A language-independent system for data-driven dependency parsing. *Natural Language Engineering*, 13(2):95–135.

Oliva, J., J. I. Serrano, M. D. Del Castillo, y A. Igesias. 2013. A SMS normalization system integrating multiple grammatical resources. *Natural Language Engineering*, 19:121–141, 0.

Pak, A. y P. Paroubek. 2010. Twitter as a corpus for sentiment analysis and opinion mining. En *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*, Valletta, Malta, Mayo. European Language Resources Association (ELRA).

Pang, B., L. Lee, y S. Vaithyanathan. 2002. Thumbs up? sentiment classification using machine learning techniques. En *Proceedings of EMNLP*, páginas 79–86.

Platt, J. C. 1999. Advances in kernel methods. MIT Press, Cambridge, MA, USA, capítulo Fast training of support vector machines using sequential minimal optimization, páginas 185–208.

Saralegi Urizar, X. y I. San Vicente Roncal. 2012. Detecting Sentiments in Spanish Tweets. En *TASS 2012 Working Notes*, Castellón, Spain.

Spencer, J. y G. Uchyigit. 2012. Sentimentor: Sentiment Analysis on Twitter Data. En *The 1st International Workshop on Sentiment Discovery from Affective Data*, Bristol, United Kingdom.

Taboada, M., J. Brooke, M. Tofiloski, K. Voll, y M. Stede. 2011. Lexicon-based methods for sentiment analysis. *Computational Linguistics*, 37(2):267–307.

Thelwall, M., K. Buckley, y G. Paltoglou. 2011. Sentiment in Twitter events. *J. Am. Soc. Inf. Sci. Technol.*, 62(2):406–418.

Thelwall, M., K. Buckley, y G. Paltoglou. 2012. Sentiment strength detection for the social web. *J. Am. Soc. Inf. Sci. Technol.*, 63(1):163–173.

Thelwall, M., K. Buckley, G. Paltoglou, D. Cai, y A. Kappas. 2010. Sentiment in short strength detection informal text. *J. Am. Soc. Inf. Sci. Technol.*, 61(12):2544–2558.

Trilla, A. y F. Alías. 2012. Sentiment Analysis of Twitter messages based on Multinomial Naive Bayes. En *TASS 2012 Working Notes*, Castellón, Spain.

Turney, P. D. 2002. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. En *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ACL '02, páginas 417–424, Stroudsburg, PA, USA. ACL.

Vilares, D., M. A. Alonso, y C. Gómez-Rodríguez. 2013. Clasificación de polaridad en textos con opiniones en español mediante análisis sintáctico de dependencias. *Procesamiento de Lenguaje Natural*, 50:13–20.

Villena-Román, J., S. Lana-Serrano, J. C. González Cristóbal, y E. Martínez-Cámara. 2013. TASS - Worshop on Sentiment Analysis at SEPLN. *Procesamiento de Lenguaje Natural*, 50:37–44.

Zhang, Lei, Riddhiman Ghosh, Mohamed Dekhil, Meichun Hsu, y Bing Liu. 2011. Combining lexicon-based and learning-based methods for Twitter sentiment analysis. Informe Técnico HPL-2011-89, HP Laboratories, Palo Alto, CA.